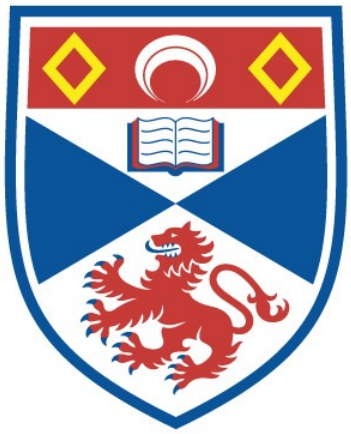
ETHNOGRAPHICALLY-INFORMED DISTRIBUTED PARTICIPATORY DESIGN FRAMEWORK FOR SOCIOTECHNICAL CHANGE: CO-DESIGNING A COLLABORATIVE TRAINING TOOL TO SUPPORT REAL-TIME COLLABORATIVE WRITING

Abd Alsattar Ardati

A Thesis Submitted for the Degree of PhD
at the
University of St Andrews



2023

Ethnographically-informed Distributed Participatory Design Framework for Sociotechnical Change: Co-Designing a Collaborative Training Tool to Support Real-time Collaborative Writing

Abd Alsattar Ardati

University of
St Andrews

This thesis is submitted in partial fulfilment for the degree of

Doctor of Philosophy (PhD)

at the University of St Andrews

March 2023

Abstract

Although Wikipedia's immense success is partially due to its support of the asynchronous collaboration model, researchers argue that the bureaucratic rules and technical infrastructure enabling it feed into Wikipedia's content bias. Attempts to introduce different collaboration models have so far failed, but the fact that they have occurred persistently over time suggests that at least part of the Wikipedia community favours incorporating features such as real-time collaborative editing.

My research is founded on the argument that the advantageous aspects of the asynchronous model should be preserved, although the existing model needs to be complemented by real-time collaboration in settings such as Wikipedia training events. This thesis describes a Participatory Design process resulting in a prototype called WikiSync, a system that introduces real-time collaboration for the Wikipedia community using a responsible design approach that is respectful of Wikipedia's rich social structure and history.

Furthermore, my research has produced an adaptive methodology for co-designing sociotechnical solutions in a geographically distributed community. After an in-depth observation of online Wikipedia training and the existing community innovation processes, my participatory design sessions have helped create a mutual learning environment for co-designing WikiSync in tandem with the community, while addressing a wide range of their concerns about real-time collaboration. I also consulted the broader Wikipedia community using an online social ideation and voting tool to evaluate the desirability and applicability of the solution. Finally, the resulting ethnographically-informed distributed Participatory Design framework provides an innovation process for involving a diverse, widely distributed online community in co-designing sociotechnical solutions.

To my parents,

Hayat Arnaout

&

Taiseer Ardati,

whose belief in me was my guiding light.

Acknowledgements

My mother, Hayat, father, Taiseer, and siblings, Mostafa, Nabil, Amina, Samah, and Areej, mere words cannot express my gratitude for your unconditional love and support; thank you from the bottom of my heart. I am also especially grateful to my faithful, patient partner, Hiba, for going the extra mile and being the wonderful person you are.

I wish to express my sincere appreciation and gratitude to my supervisors, Dr Alexander Voss, Dr Uta Hinrichs, Professor Ian Gent, and Dr Angela Miguel, whose continuous invaluable guidance and advice have made my research journey an inspiring experience and helped me develop academically as well as personally.

Thank you, Alex, for being the great human being you are. My life would have been significantly different had our paths not crossed in 2018.

It is difficult for me to put into words my gratitude for friends who have lent me their generous support, and whom I cannot thank enough. Tomas Vancisin, Xu Zhu, Daniel Koudouna, Mustafa Abdelwahed, Anas Arnaout, Sulyman Hammoud, Gemma Marina, Kaixuan Wang, Janice Deary, Kirsty Ross, Sara Thomas, Fadi Abdelhak, Marjorie Gourlay, and the Global Share's team. You have been more than just friends; you are my second family whose camaraderie is the safety net I have leaned on multiple times on this long PhD journey. Your encouragement, small chats, and comments on my work have helped me see the finish line.

I also extend my thanks to the School of Computer Science, the University of St Andrews, St Leonard's Postgraduate College, Wikimedia UK and the IDEA Network for their vital support, and for giving me this opportunity to research something close to my heart.

I am indeed lucky to have chosen such a career path that has led me to be surrounded by such people as my research collaborators, Kirsty Ross, Sara Thomas, Richard Nevell and Dariusz Jemielniak, who have joined me on this research journey and offered considerable support.

Finally, I would like to express my sincere gratitude to my research participants and the Wikipedia community for generously supporting this research. "Alone we can do so little; together we can do so much" ― Helen Keller

Funding

This work was supported by the University of St Andrews (St Leonard's Postgraduate College and School of Computer Science) through St Leonard's Scholarship.

Declarations

Candidate's declaration

I, Abd Alsattar Ardati, do hereby certify that this thesis, submitted for the degree of PhD, which is approximately 83,000 words in length, has been written by me, and that it is the record of work carried out by me, or principally by myself in collaboration with others as acknowledged, and that it has not been submitted in any previous application for any degree. I confirm that any appendices included in my thesis contain only material permitted by the 'Assessment of Postgraduate Research Students' policy.

I was admitted as a research student at the University of St Andrews in May 2019.

I received funding from an organisation or institution and have acknowledged the funder(s) in the full text of my thesis.

Date: 24/08/2023

Supervisor's declaration

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of PhD in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree. I confirm that any appendices included in the thesis contain only material permitted by the 'Assessment of Postgraduate Research Students' policy.

Date: 24/08/2023

Permission for publication

In submitting this thesis to the University of St Andrews we understand that we are giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. We also understand, unless exempt by an award of an embargo as requested below, that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that this thesis will be electronically accessible for personal or research use and that the library has the right to migrate this thesis into new electronic forms as required to ensure continued access to the thesis.

I, Abd Alsattar Ardati, confirm that my thesis does not contain any third-party material that requires copyright clearance.

The following is an agreed request by candidate and supervisor regarding the publication of this thesis:

Printed copy

No embargo on print copy.

Electronic copy

No embargo on electronic copy.

Date: 24/08/2023

Date 24/08/2023

Underpinning Research Data or Digital Outputs

Candidate's declaration

I, Abd Alsattar Ardati, hereby certify that no requirements to deposit original research data or digital outputs apply to this thesis and that, where appropriate, secondary data used have been referenced in the full text of my thesis.

Date 24/08/2023

Table of Contents

List of Figures

List of Tables

Glossary

Literature uses such terms as Participatory Design and Co-design interchangeably. Therefore, the key terms used in this thesis are clarified below. Some of these definitions are adapted from Sanders et al. (2010) and Constantin et al. (2021):

Participatory Design (PD) is used to describe the approach, philosophy, or methodology that underpins the choices in this research. While recently, the political aspect of participation or empowerment of a democratic decision-making process was weakened in favour of the focus on user participation, the proposed framework of this thesis emphasises the importance of a strong presence of user empowerment at the very early stage of the problem definition and through the process, especially for the marginalised voices.

Co-Design is used as a term to describe the activities, actions, methods and techniques guided by the PD process and principles.

Distributed Participatory Design (DPD) describes participatory design projects where the co-design activities involve geographically distributed participants, leading to coordinating these activities to mitigate the adverse effects of not having the team members interact in person. While most research, including mine, refers to online PD as the primary way to facilitate DPD, such coordination for participants' contributions can occur offline.

Tools can be defined as means to conduct and facilitate a co-design activity, such as virtual collaboration boards or video conferencing tools. Another example of these tools is the material used on the virtual boards, such as a sticker board that shows tools used by the community on these virtual boards.

Activity is a stage in a PD process or session; this stage may involve the use of a method or a tool to achieve the goal of the participatory design session phase.

Techniques are the way tools are used to achieve the goal of the activity. For example, combining similar ideas or categorising and linking sticky notes.

Method is used to describe a group of tools and techniques utilised to achieve a certain goal. For instance, observation may involve the utilisation of several tools and techniques to bring insights into community practices.

Chapter 1 Introduction

Wikipedia is a widely popular online encyclopaedia created and sustained by volunteers. However, it faces many challenges, such as content bias (Suh et al., 2009; Lam et al., 2011; Geiger & Halfaker, 2013; Nakamura et al., 2013; Hube, 2017; TeBlunthuis et al., 2018; Yun et al., 2019; Oeberst et al., 2020; Menking & Rosenberg, 2021). Specifically, researchers have found that the current asynchronous collaborative model is prone to gender bias (Collier & Bear, 2012; Bear & Collier, 2016), "collective biases" (Oeberst et al., 2016), and hindsight bias (Meuer et al., 2022). On the other hand, researchers such as Shane-Simpson & Gillespie-Lynch (2017) and Langrock & González-Bailón (2022) have shown how the potential of new models of collaboration, feedback, and Wikipedia training may address bias. For example, Shane-Simpson & Gillespie-Lynch (2017, p. 324) found that "the presence of visible female editors on Wikipedia and broader encouragement of the use of constructive feedback may help to alleviate some of the current Wikipedia gender gap."

However, efforts to introduce change to Wikipedia collaboration models have a high probability of coming to a halt (*Wikipedia Flow Project Talk Page*, 2016; *Real Time Group Editing Proposal - Community Wishlist Survey*, 2016; *Real-Time Collaboration Timeline - MediaWiki*, 2022; *Wikimedia's Phabricator*, 2022).[1] Most changes have not "gained wide acceptance" or have been "heavily criticised" by the community (*Talk Pages Consultation 2019 - MediaWiki*, 2019). It is challenging for the Wikimedia Foundation,[2] the organisation that supports the Wikipedia infrastructure, to understand and address these complex social and technical challenges, such as the case of introducing real-time collaboration to Wikipedia (Ananian et al., 2018).

Weak knowledge transfer and communication channels between the Wikipedia community and Wikimedia pose a considerable threat to Wikipedia. For example, in 2002 significant changes in the platform led to the migration of most Spanish Wikipedia active contributors out of the platform over a controversy regarding adding ads to Wikipedia ('List of Wikipedia Controversies', 2022). Another incident of a "community revolt" over working on a new search engine project resulted in the Wikimedia CEO's resignation (Hern, 2016). Therefore, research to find a new collaborative innovation process that brings the community and Wikimedia closer in designing solutions is needed.

[1] Italic references means that the source is a webpage, such as a Wikipedia page.
[2] https://wikimediafoundation.org/

My thesis focuses on how we design solutions serving peer production online communities. It is focused on creating a shared online space for designers and the community to make a positive and sustainable change by mapping the challenges that hinder change to Wikipedia's collaborative practices.

In my thesis, I offer a prototype for a tool called WikiSync, the first Wikipedia training tool that involves real-time collaborative editing of Wikipedia articles co-designed using a distributed approach that involves the Wikipedia community through several phases that vary in focus and scope of user participation.

1.1 Background

Arguably the most successful socio-technical knowledge infrastructure built through volunteer effort (Geiger & Halfaker, 2013), Wikipedia expands the boundaries of open knowledge through its breadth, depth, and scale of participation (Cummings, 2020). Since its inception in 2001, this online encyclopaedia has grown to be among the top ten most widely visited websites worldwide (Alexa - Top sites, 2022). Each month, Wikipedia's 300+ language versions receive billions of unique visits (Statista, 2021). Leading tech companies, such as Google and Facebook, rely on peer-produced content to verify facts on their platforms (Perez, 2020; Sullivan, 2020; Vincent & Hecht, 2020), as well as feed its content into their machine-learning algorithms (Graham et al., 2014; Singhal, 2012; Zou & Schiebinger, 2018; *Google Translate - MediaWiki*, 2022).

Wikipedia's immense success is partially due to its support of *liquid collaboration* (Jemielniak & Raburski, 2014), where editing takes place without direct interaction between collaborators, as changes are integrated through a technical, bureaucratic system rather than through direct human collaboration or negotiation between contributors. In other words, "more often in Wikipedia, the collaboration process is not synchronous interpersonal collaboration, it is asynchronous individuals writing and editing articles with a technology to mediate and aggregate the output" (Collier & Bear, 2012, p. 385). This collaboration model has arguably been a key factor in Wikipedia's emergence as a global open knowledge infrastructure (Yun et al., 2019).

Challenges of Wikipedia

Wikipedia faces challenges in terms of societal inequalities and social conflicts, leading to the production of biased information (Suh et al., 2009; Lam et al., 2011; Geiger & Halfaker, 2013; Nakamura et al., 2013; Hube, 2017; TeBlunthuis et al., 2018; Yun et al., 2019; Oeberst et al., 2020; Menking & Rosenberg, 2021), exclusion of information (Graham, 2011; Graham &

Hogan, 2014; Ford, 2017), as well as misinformation and disinformation (Kumar et al., 2016). Additionally, the English Wikipedia faces a low retention rate of new editors (Morgan & Halfaker, 2018). The Wikimedia Foundation (Wikimedia) is aware of these challenges, documenting them in their reports and actively looking for solutions (Wikimedia Foundation, 2011). If left unresolved, these challenges may threaten Wikipedia's hard-won status as a reliable source of knowledge.

My thesis focuses on the interplay between the specific social and technical challenges that the community of Wikipedia trainers and new editors face. Wikipedia, as a socio-technical infrastructure and community enterprise, suffers from a number of difficulties, including a lack of diversity in its editor community and a clear imbalance of power between established editors and newcomers or more casual contributors (Morgan & Halfaker, 2018; Musicant et al., 2011; Narayan et al., 2017; *Wikimedia Growth Team*, 2022). These inequities affect both the underlying technical infrastructure of source code and hosting arrangements and, more importantly, the development of Wikipedia content. Research shows that Wikipedia articles are governed and written by relatively few "super-editors" who are praised for their high dedication to enhancing their quality (Yun et al., 2019). On the other hand, they are critiqued for reducing the diversity of contributors, leading to the production of many articles with "biased coverage" (Yun et al., 2019). For example, examining gender bias on Wikipedia, Wagner et al. (2015) analysed biographical articles about women. Their findings show that women's articles are more frequently linked to men's than vice versa. In addition, Wagner et al. found that women's biographies cover more romantic and family-related content compared to those of men.

Wikipedia's technological infrastructure has been blamed for causing bias by demanding that editors adapt to an environment orientated toward men (Ford & Wajcman, 2017). Moreover, the technical development and innovation process still tend to be driven mainly by the vision of this "core community" of editors.

The Role of Collaboration in Wikipedia

As discussed earlier, some of these challenges are attributable to the model of collaboration, which does not mitigate persistent issues, such as systemic bias, since:

"The collaboration process is not synchronous collaboration with others on a personal basis; rather, the process entails asynchronous writing and editing of articles. A large part of contributing, then, is one person editing the work of unknown others and competing for one's words to prevail." (Bear & Collier, 2016, p. 256).

Furthermore, Ford and Wajcman (2017) argue that bureaucratic rules and the technical infrastructure that enable liquid collaboration play a major role in generating such bias. They are also a barrier to entry and create an "antagonistic" environment for new editors. (Collier & Bear, 2012; Goldman, 2009; Suh et al., 2009; Morgan et al., 2013).

Some researchers attribute such inequity to the collaboration style, i.e. liquid collaboration (Jemielniak & Raburski, 2014), bureaucratic coordination as a hindrance, and asynchronous editing, which favours the 'lone wolf' over genuinely collaborative and explicit coordination practices (Kittur et al., 2006; Callahan & Herring, 2011; Hube, 2017; Yun et al., 2019). It is noteworthy that the 'lone wolf' phenomenon is not limited to Wikipedia but is equally seen in other projects (Chełkowski et al., 2016; Markel & Guo, 2020).

Reviewing Bear & Collier's (2012, p. 385) discussion of how the asynchronous model encourages an environment of "competing (rather than collaborating) for one's words to cancel out the contributions of others", one cannot help but wonder why Wikipedia articles cannot be edited in new models that facilitate "synchronous interpersonal collaboration".

Real-time Collaboration and Wikipedia Training

Wikipedia uses Wiki software known for asynchronous collaboration (Decker et al., 2007). Introducing real-time editing for all users would have ramifications that affect the entire socio-technical system. For instance, real-time editing makes attributing multiple authors more complicated (Ananian et al., 2018). However, it is worth examining the possibility of introducing this new mode of collaboration for small groups of editors in remote Wikipedia training sessions[3] or WikiProject,[4] something already suggested by the community (*Wikimedia's Phabricator*, 2022) but not yet implemented. The fact that they have been suggested multiple times indicates the existence of arguments that favour real-time collaboration. Moreover, one of the staff members on Wikimedia's Phabricator page who discussed the feature suggested that "real-time editing might be useful for Edit-a-Thons and student groups." (*Wikimedia's Phabricator*, 2022).

Since the COVID-19 outbreak, the remote model of collaboration has become prevalent, and video conferencing and remote collaboration tools have suddenly become the main way to accommodate the need to take these activities online. To help accomplish this task, academic institutions and professional organisations have adopted a wide range of tools to meet their

[3] https://wikimedia.org.uk/wiki/Training
[4] A team of Wikipedia editors who collaborate to enhance a certain area on Wikipedia https://en.wikipedia.org/wiki/Wikipedia:WikiProject

online transition needs through technologies that support real-time communication, such as Zoom,[5] WebEx,[6] and Microsoft Teams,[7] and collaborative writing tools, including Google Docs[8] and Microsoft 365.[9] This has led Markel & Guo (2020) to call on designers to develop solutions that provide learners with a more inclusive and dynamic learning environment that rivals the face-to-face one. This thesis is a response to this call.

Several modalities of Wikipedia editing are discussed in research (cf. (Kittur et al., 2006; Callahan & Herring, 2011; Hube, 2017; Yun et al., 2019)) and described in this thesis. Instead of the 'lone wolf' type mentioned earlier, Wikimedia UK specifically aims to foster a collaborative learning environment to allow newcomers to move from following instructions and community rules to working in teams and learning from each other. Wikimedia UK ceased its face-to-face training model in March 2020 due to the COVID-19 pandemic (Thomas, 2021). Consequently, the help of Wikimedia experts and the wider community was sought to find an online set of tools as an alternative for onboarding new Wikipedia editors. As a result, in June 2020, Wikimedia UK ran its first online 'Training the Trainers' programme, which consisted of three sessions on designing and developing online training programmes and using the tools at one's disposal, such as Zoom, Are.na,[10] and Etherpad.[11] According to Wikimedia UK, the online mode of training is here to stay (Thomas, 2021; *Online Train the Trainer 2022 - Call for Participants*, 2021). This presents an opportunity to make training available regardless of geographic location, also including new Wikipedia editors who cannot participate in face-to-face meetings for some reason (ibid.).

Nevertheless, conducting training is challenged by Wikipedia's technology, which does not allow for real-time collaboration and coordination. Sliger Krause et al., (2017) found that:

> "Participants could not all effectively participate simultaneously since we observed that as the number of people working on the article at the same time increased, it became much more difficult for them to avoid crossing digital paths in the Wikipedia platform. This resulted in "editconflicts," in which two authors have submitted different and overlapping changes to the same section [...] resolving these conflicts generally requires more advanced

[5] https://zoom.us/
[6] https://www.webex.com/
[7] https://www.microsoft.com/en-gb/microsoft-teams/group-chat-software
[8] https://docs.google.com
[9] https://www.office.com/
[10] https://www.are.na/
[11] https://etherpad.org/

> familiarity with the Wikipedia interface and markup" (Sliger Krause et al., 2017, p. 17).

Newcomers find Wikipedia stripped of the familiar tools that already support real-time collaborative writing with no edit conflicts, such as Google Docs, that incorporate these very features that facilitate real-time collaborative processes.

During Wikipedia training, newcomers are often encouraged to collaborate on tasks such as gathering information for articles with others using real-time collaborative editing tools such as Etherpad[12] (*Syberthon Edit-a-Thon*, 2020; *UNESCO Wiki4Women Edit-a-Thon (India)*, 2020) or Google Sheets,[13] and ask for help using real-time chat tools, such as the one embedded in a training conference call, or through Wikipedia's Internet Relay Chat (IRC).[14] However, once the training sessions are over, the newcomers face a very different style of collaboration, which Jemielniak and Raburski refer to as the liquid collaboration, characterised not by real-time forms of collaboration but by a model of tying together asynchronous contributions from individuals via well-established bureaucratic processes and social norms, for example, guidelines on how to avoid "edit conflict".[15] As Jemielniak and Raburski put it, "edits contribute to the development of the articles because they are ordered and aggregated by the common technological and institutional framework, rather than thanks to any collaborative efforts" (Jemielniak & Raburski, 2014, p. 96). Put succinctly, if somewhat crudely: Wikipedia is more an outcome of crowdsourcing and bureaucratic processes than collaboration, per se.

The need for training Wikipedia newcomers to use online tools instead of face-to-face interactions has inspired the direction of my PhD, where the challenges faced by Wikipedia as a socio-technical project are examined from the perspective of collaboration. This thesis discusses and synthesises different areas of research, highlighting opportunities for introducing real-time collaboration for newcomers who attend online training in a way that still respects Wikipedia's rich social structure and history. This led to the design of WikiSync, a Wikipedia training system that addresses how to introduce real-time collaboration for Wikipedia.

Improving the onboarding process has proved to significantly improve the retention of new editors, which is one of the main challenges for Wikipedia (Morgan & Halfaker, 2018). In collaboration with Wikimedia UK, my research examined Wikipedia's remote training and editing experience to inform the co-design of an online tool that can support Wikipedia trainers'

[12] https://wikitech.wikimedia.org/wiki/Etherpad.wikimedia.org
[13] https://www.google.co.uk/sheets/about/
[14] https://en.wikipedia.org/wiki/Wikipedia:IRC
[15] https://en.wikipedia.org/wiki/Help:Edit_conflict

efforts to train groups of newcomers remotely. The subsequent co-design with trainers and trainees resulted in a design for WikiSync that supports a new way of real-time collaboration in the tailoring of Wikipedia training sessions in addition to editing Wikipedia articles. Finally, a prototype of WikiSync was reviewed with the community to facilitate discussions on its future implementation and to investigate sustainability options.

1.2 Problem Statement & Research Questions

Procedures, guidelines, and biases lead to a steep learning curve for becoming an active Wikipedia editor and how to write content "the Wikipedia way." (Goldman, 2009; Suh et al., 2009; Reboot, 2017, p. 22). This is exacerbated by recognised usability issues with Wikipedia's two different editing interfaces (visual and source editors)[16] and existing commitments to specific technologies, collaboration, and innovation models (Ford & Wajcman, 2017; Reboot, 2017; Gluza et al., 2021). Despite the Wikipedia Foundation's efforts to improve the onboarding experience, many new editors do not return to make subsequent edits (TeBlunthuis et al., 2018).

Even though Wikipedia is known to be made possible by collaboration, the asynchronous editing and communication model is not intuitive for those with no "coding or programming knowledge [is] not intuitive or discoverable, making it difficult for new editors to learn and progress." (Reboot, 2017, p. 24). The *liquid collaboration* model (Jemielniak & Raburski, 2014) of *asynchronous* editing reconciled through a bureaucratic process does not sit well with prevalent user expectations (Gluza et al., 2021) and is one of the "new editors' greatest challenges" (Reboot, 2017, p. 22).

Real-time collaboration has been requested by the Wikipedia community and proposed by Wikimedia staff as a viable option to explore, but the obstacles to its introduction are significant given the existing collaboration model that made Wikipedia successful in the first instance (*Wikimedia's Phabricator*, 2022; *Real-Time Collaboration Timeline - MediaWiki*, 2022). As Ananian et al. (2018) highlighted at three different Wikimania conference[17] presentations between 2015 and 2018: "Lots of talk among developers about technical mechanisms for real-time collaboration. Very little talk about the social aspects or impact on community. Not much dialog between developers and editors and readers. Let's start to fix that!"

[16] "Wikipedia gives you two options for editing. The Visual Editor lets you edit the text on the screen just as it appears on a Wikipedia page, similar to other word processing applications and websites. Source editing lets you see the underlying Wikicode, which resembles HTML." https://dashboard.wikiedu.org/training/students/how-to-edit/visual-editing-vs-source-editing-v2

[17] https://en.wikipedia.org/wiki/Wikimania

My PhD research investigates the challenges of Wikipedia's collaborative innovation using the case of introducing alternative models of collaboration. It focuses on addressing the following main research question:

How can we involve Wikipedia's community and newcomers in designing a solution for introducing real-time collaboration while at the same time being responsible and respectful of Wikipedia's rich social structure and history?

The research's main question is divided into sub-questions that will be addressed throughout the thesis chapters:

Q1. What are the challenges of introducing real-time collaboration to the Wikipedia community, especially in training events?

Q2. What are the key features of a real-time collaborative tool for Wikipedia newcomers?

Q3. How can we co-design a tool and detailed training scenarios that would benefit from real-time collaborative editing?

Q4. How can we apply and possibly adapt participatory design processes for online communities and Wikipedia specifically?

1.3 Primary Contributions

The following are the main contributions of my PhD research:

C1: Exploration of the potential of introducing new real-time writing technology in Wikipedia training events.

My research provides a synthesis of related research work discussing the benefits of real-time collaborative editing and existing guidelines for collaborative editing tools outside the Wikipedia context, providing lessons learned that can help Wikimedia and the Wikipedia community. It also presents a picture of a large portion of Wikipedia's innovation process and its strengths and weaknesses, demonstrating a gap in research that considers the issue of introducing real-time collaboration and training for Wikipedia.

The thesis covers research on Wiki editing in educational contexts and general design guidelines for collaborative editing tools. It outlines pathways to new technologies and practices for Wikipedia trainers and editors as well as those who are researching solutions to the problem of bias in open knowledge, collaborative group Wikipedia article writing, and the design of crowdsourcing systems.

C2: Analysis of research findings on the evolving *liquid collaboration* model in Wikipedia and its links to bias in the content.

My thesis provides an analysis of the evolving liquid collaboration model and several coordination models, supporting reasons for combining real-time and asynchronous collaborative writing, which can help Wikipedia editors in situations such as training sessions, despite the challenges that must be overcome to facilitate the novel collaborative approach. Moreover, it lays the foundations for future research aimed at investigating Wikipedia training group dynamics and the impact of technological choices on issues such as systematic bias.

C3: Providing a prototype of a possible solution to introduce real-time collaboration through Wikipedia training.

My PhD research offers a prototype co-designed with the Wikipedia community that can be a stepping stone for introducing an alternative real-time collaboration technology to Wikipedia while at the same time being responsible and respectful of its rich social structure and history.

C4: An adaptive framework for co-design solutions for small research groups in a distributed setting.

My research provides a distributed framework for co-design solutions for small research groups in a distributed setting. It is useful for teams interested in developing conceptual work and designing innovation processes that systematically co-evolve computer-supported collaborative systems and overcome socio-technical issues, especially in the education context that involves learning through collaborative writing.

C5: The application of the distributed co-design framework to Wikipedia's innovation process.

My PhD research offers a new way to involve Wikipedia's community and newcomers in designing solutions, such as introducing major features to MediaWiki.

1.4 Thesis Structure

Chapter 2 provides a background to Wikipedia's main challenges regarding retaining new editors and the lack of editor diversity contributing to biased content. In this chapter, I define concepts and technologies, such as the Wiki software, outlining research on collaborative Wikipedia editing infrastructure. The chapter covers the interdependency between Wikipedia's social and technical elements and its design approaches. It then covers the synthesis of the benefits of supporting real-time collaborative writing.

Chapter 3 describes and justifies the overall methodological approach of my PhD research in the context of Wikipedia onboarding, training and collaboration. It discusses the co-design literature, elaborating on the importance of the direct, democratic, and diverse involvement of those affected by design in shaping the solution outcome, thereby ensuring its sustainability.

Chapter 4 describes the design and findings from an observational study that explores the Wikipedia innovation process and analyses observations of Wikipedia training to shape a series of Participatory Design (PD) sessions for the following three chapters.

Chapters 5, 6 and 7 describe the methods, procedures, and findings from a series of nine PD sessions focused on prototyping WikiSync and developing a framework for designing solutions serving peer production online communities. The sessions are divided into 3 phases: the Discovery phase (2 sessions), the Conceptualisation phase (4 sessions), and the Design phase (3 sessions).

Chapter 5 describes the design and findings from co-design sessions conducted with new and established Wikipedia community members focused on the Discovery phase. In these sessions, we collaborated on defining the current training activities flow. The chapter also outlines how the sessions in the Discovery phase helped give insights into the challenges, opportunities and potentials of a new collaboration model in training.

Chapter 6 discusses the Conceptualisation phase. It focuses on defining the user groups who can benefit from the new collaboration model. It covers the collaboration process on a new system design concept, usage scenario and guidelines. It also includes a proposal for new real-time collaboration models using WikiSync. This phase involved taking a step back to look critically at the design direction with new participants from more countries.

Chapter 7 covers the Design phase. It demonstrates how emerging themes from the Conceptualisation phase of the PD sessions help map the activities of a new training model backed with a storyboard and sketches. In addition, this phase involved the broader community in evaluating the design concept's desirability, applicability and inclusivity using an online social ideation and voting tool. Finally, after developing the final prototype for this research project, the participation was narrowed down for a final review and planning session for sustaining the efforts in developing and implementing the system for the future after this PhD research.

Chapter 8 outlines my main contributions to Wikipedia and the design research community, proposing a co-design approach that can be utilised with limited resources for an inclusive innovation process.

Chapter 2 Wikipedia and Real-time Collaboration

Many social and technical factors hinder change in Wikipedia as it strives to meet the evolving needs of its global community of editors and readers. The lack of agility can be partially comprehended by reviewing the conflicting perspectives behind the scenes of Wikimedia's celebration of its "Fresh New Look" (*Wikipedia Gets a Fresh New Look*, 2023), a much-needed refresh of the website design that was a decade in the making. In this thesis, I ask if Wikipedia can afford such lag in evolving its collaboration models. The delay in changing Wikipedia's UI shows that addressing this research project question, which targets core functionality in how these platforms operate, is compounded by several socio-technical issues, one of which is a lack of research in this area.

At the time of writing, no literature exists on the potential of introducing new, real-time writing to Wikipedia. Moreover, Wikipedia training holds "the potential to be part of solving these [content biases and gaps and retaining new editors] problems. [And] despite being around for many years now, systematic investigations of how editors experience these collaborative writing events are missing, to a large degree." (Gluza et al., 2021, p. 10). Therefore, the literature review in this chapter covers research from several domains that, taken together, can provide insights that would lead to answering the research questions.

This chapter aims to document this literature review process and provide an interdisciplinary synthesis of research on Wikipedia and Wiki software more generally. It covers different collaborative models' strengths and interrelated challenges, demonstrating the potential of introducing new real-time writing technology in Wikipedia training events.

Furthermore, in this chapter, I argue that the research community should explore a novel innovation process to help the Wikipedia community identify the benefits of new collaboration modes on their platform, such as synchronous collaboration. The introduction of new forms of collaborative training, which are forced by COVID-19 measures, may encourage one to revisit the question of whether Wikipedia should adopt new collaborative models.

Implementing new collaboration modes, such as real-time collaboration in Wikipedia, is now possible from a technical point of view (*Waiting for Real-Time Collaboration*, 2017). However, the true challenge lies in the social impact these would have on community norms and rules, which have been built around and support Wikipedia's asynchronous editing model "by anyone, at any time" (*Wikipedia:FAQ/Contributing - Wikipedia*, 2021). Challenges, such as author attribution, real-time vandalism and harassment, and the co-existence of real-time

and non-real-time edits, have been discussed widely at the official annual conferences of the Wikimedia Foundation (Ananian et al., 2018).

This leads me to argue that rather than replace the current model with synchronous collaborative writing, specific usage scenarios should be considered. Examples might include Wikipedia training sessions for new editors, Edit-a-Thons, and Wikipedia projects. These scenarios are marked by the simultaneous, explicit task coordination of participants who usually share a common sense of purpose and are physically collocated or connected via a synchronous communication channel. Often, they also have other things in common, such as cultural background or common concerns.

The coordination style plays a major role in the Wikipedia community; researchers such as Kittur and Kraut (2008) have demonstrated that, when coordination is implicit, higher quality articles often result from having a larger rather than a small number of editors undertaking the majority of work, some of them set the direction of the article, while other editors would develop the rest of the article gradually. However, higher-quality, explicitly coordinated articles correspond with fewer editors in the "formative stages, when its structure is highly unconstrained" (Kittur & Kraut, 2008, p. 9). Since training events help provide the environment to facilitate collaboration, where a small subset of editors set the direction of the article and lay its foundation early on for later contributions by others, I argue that dividing the trainees into small teams to coordinate the writing task explicitly could improve the resulting quality of the Wikipedia article, though the need to provide this small team with the synchronous collaborative writing model inside Wikipedia remains to this day unaddressed.

In Wikipedia training, newcomers are often advised against working together simultaneously so as to avoid possible conflict and friction (Sliger Krause et al., 2017). In 2001, at Wikipedia's launch, it may have made sense to warn new users against editing a document in real-time. Such advice was supported by research twenty years ago, when Noël & Robert's (2003, 2004) research showed that group members preferred to work independently on collaborative documents in a "hands-off" style over using available collaborative writing applications. However, more recent research shows that this is no longer the case. Using the DocViz[18] tool for visualising the contribution of real-time writing, Olson et al. (2017) analysed 96 Google Docs collaboratively written as undergraduate assignments by 32 teams of four editors. Their research challenges the above-mentioned results, as student groups preferred working in real-time on collaborative writing tasks in their experiments. In fact, some of the

[18] https://chrome.google.com/webstore/detail/docuviz/hbpgkphoidndelcmmiihlmjnnogcnigi?hl=en

groups only contributed to the assignment when the whole team gathered to work simultaneously.

With the prevalence of advanced collaborative writing tools in today's technological landscape, such as Google Docs, being told to write a Wikipedia article only using the asynchronous model might discourage new volunteer editors who are accustomed to simultaneous writing. If the new editors overcome the early challenges of the asynchronous editing model, in the long run, the existing Wikipedia model is expected to work out fine for them, although not through short-term collaboration efforts. Moreover, a new collaborative writing model could be offered to people receiving training to edit Wikipedia in group settings.

This chapter explores examples of these new models of collaboration that could be useful in opening up further discussions about Wikipedia's default model of collaboration, which is the premise of Chapters 4 to 8 that describe the participatory design of a tool for synchronous collaboration.

Such new collaboration models open new prospects of contributing to addressing Wikipedia's challenges, such as the knowledge gaps discussed in 2.1. For example, page translations are best done by small teams of people coordinating their demanding translation tasks explicitly and working synchronously, perhaps supported by tools other than the Wikipedia Talk pages,[19] such as an audio channel. I would argue that multiple areas exist where it is not only possible but also more efficient to organise short-focused collaborations rather than rely on the liquid collaboration model characterised by the indirect coordination of tasks.

It makes sense that supporting groups of contributors aiming to create new pages would be more effective, helping Wikipedia reduce the number of its stubs, which constitute 37% of the English Wikipedia and 78% of its Arabic version. Furthermore, such support would contribute to new editors having good initial experiences and being more likely to maintain their engagement with the online encyclopaedia.

Wikipedia's current collaboration model is prone to gender bias (Collier & Bear, 2012; Bear & Collier, 2016). Introducing new collaboration models may address some of the criticism levelled at Wikipedia on such topics as content bias. Moreover, having several contribution models that suit editors with different preferences and backgrounds leads to a more diverse community of editors contributing to Wikipedia's articles, resulting "in higher quality and less biased articles" (Kittur & Kraut, 2008, p. 1). Therefore, a sensible switch to an innovative online training delivery model that extends the explicit coordination of tasks with synchronous

[19] Pages linked with the main article pages where Wikipedians can discuss improvements to articles or other Wikipedia pages: https://en.wikipedia.org/wiki/Help:Talk_pages

collaborative writing should be embraced out of necessity and as an opportunity to make greater progress and address justified critiques.

This chapter comprises four primary sections, some of which have several subsections. Section 2.1 explores the literature on some of Wikipedia's socio-technical barriers contributing to Wikipedia's content bias and becoming an editor. In Section 2.2, connections from these barriers to the current collaboration model are examined and demonstrated, showing the importance of considering the problem from a socio-technical perspective. Section 2.3 tells Wikipedia's history with real-time collaboration, starting with technical infrastructure and ending with numerous attempts to incorporate real-time collaboration features. Taking an interdisciplinary approach, I work out lessons learned from previous innovative attempts and the use of wiki software in education settings. The broad exploration of different research areas in this chapter has helped guide the effort to address the main research question, which provides guidance that has shaped the sub-research questions that will be discussed in Chapters 4 to 7.

2.1 Challenges

The increased demand for Wikipedia content has not been met with the same level of support for its online community of volunteers who enrich and maintain Wikipedia's content (Vincent et al., 2018). Wikipedia has challenges, biases, gaps, and experienced a decline (*Wikimedia Statistics*, 2022) in the number of active editors. As discussed below, researchers have experimented with different approaches to addressing the challenges Wikipedia faces, and as discussed below, they offered recommendations concerning its social culture and technical infrastructure.

The first section covers the issue of gender bias and its socio-technical roots, while the second illustrates the severity of Wikipedia's knowledge gaps and the need to invest in new collaborative solutions that help to recruit and retain newcomers, leading to the last section, which focuses on retaining editors.

2.1.1 Gender Bias

A primary challenge for Wikipedia is its systematic gender bias in community representation and content coverage (Wagner, Garcia, et al., 2015; *Wikipedia*, 2022). Researchers and the Wikimedia Foundation, the organisation that provides the online encyclopaedia with its essential infrastructure, have offered several social and technical explanations for this gap (Lam et al., 2011; Wikimedia Foundation, 2011; Jemielniak, 2016; Ford & Wajcman, 2017). A 2011 study showed that only 9% of all Wikipedia editors were female. To respond to this inequality, the Wikimedia Foundation pledged to increase women's editorship representation to 25% by

2015 (*Gender Bias on Wikipedia - Wikipedia*, 2023). However, according to Wikipedia's co-founder Jimmy Wales, the efforts to reach the 25% mark by 2015 have failed, and a more recent survey from 2017 to 2018 to assess Wikipedia's participation diversity showed a minor decrease in women's representation, with only 8.8% of all Wikimedia projects being led by women (*Community Insights/2018 Report/Contributors - Meta*, 2023).

Jemielniak (2016) discusses the need to avoid stereotypical explanations for such a complex problem, such as the notion that women are simply not interested in Wikipedia or lack skills. One problem that occurs regularly is that new editors are not sufficiently aware of the community rules. He argues that they should be equipped with enough knowledge about editing conventions to keep editing in a conflict-prone environment.

Shane-Simpson and Gillespie-Lynch (2017) model Wikipedia editing activities at a small scale to study the factors underlying this gap by examining the collaborative editing behaviour of 203 undergraduate students who took part in this study in exchange for academic credits. The participants, who had diverse ethnic and economic backgrounds, were asked to complete a half-written article about cyberbullying. The results suggest that the high proportion of anonymous editing, which does not reveal the editors' gender, leads to fewer contributions by female editors and widens the gap "by creating the illusion (or perhaps reality) of a male-dominated and overly critical editing environment" (Shane-Simpson & Gillespie-Lynch, 2017, p. 324). The authors call for further mixed-gender collaborative editing group research, which would be needed to adequately understand the influence of collaboration patterns, as well as the style of group communication, on the degree of the gender gap. Similarly, Sydow et al. (2017) report research findings that show that such attributes as the diversity of interests and expertise in a Wikipedia editing team can boost the teamwork quality more than the size of the group, productivity, or experience of a contributor.

Ford and Wajcman (2017) suggest that the gender gap has two leading causes: Wikipedia's social culture, which is "governed by historically conservative (male) scientific understanding of expertise and authority" (Ford and Wajcman, 2017, p. 516), and Wikipedia's technological infrastructure, which requires the technical know-how that is inherently biased toward males. Eckert and Steiner (2013) argue that the "geek heritage" of Wikipedia derived from the open-source Wiki infrastructure has adversely affected its inclusivity, even though Wilson's (2000) research shows that women are more inclined to volunteer their leisure time than men.

2.1.2 Knowledge Gaps

The United Kingdom, Italy, France, Germany, and the United States account for almost 57.6% of Wikipedia edits, while Australia has more editors than all of Africa (Graham et al., 2014, 2015). France has produced twice as much content contribution as the African continent (Graham, 2014). Graham and Hogan (2014) found uneven coverage in articles about places (cities, towns, historical monuments, etc.), as can be seen in Figure 2.1, as 84% of Wikipedia articles concentrate on North America and Europe.

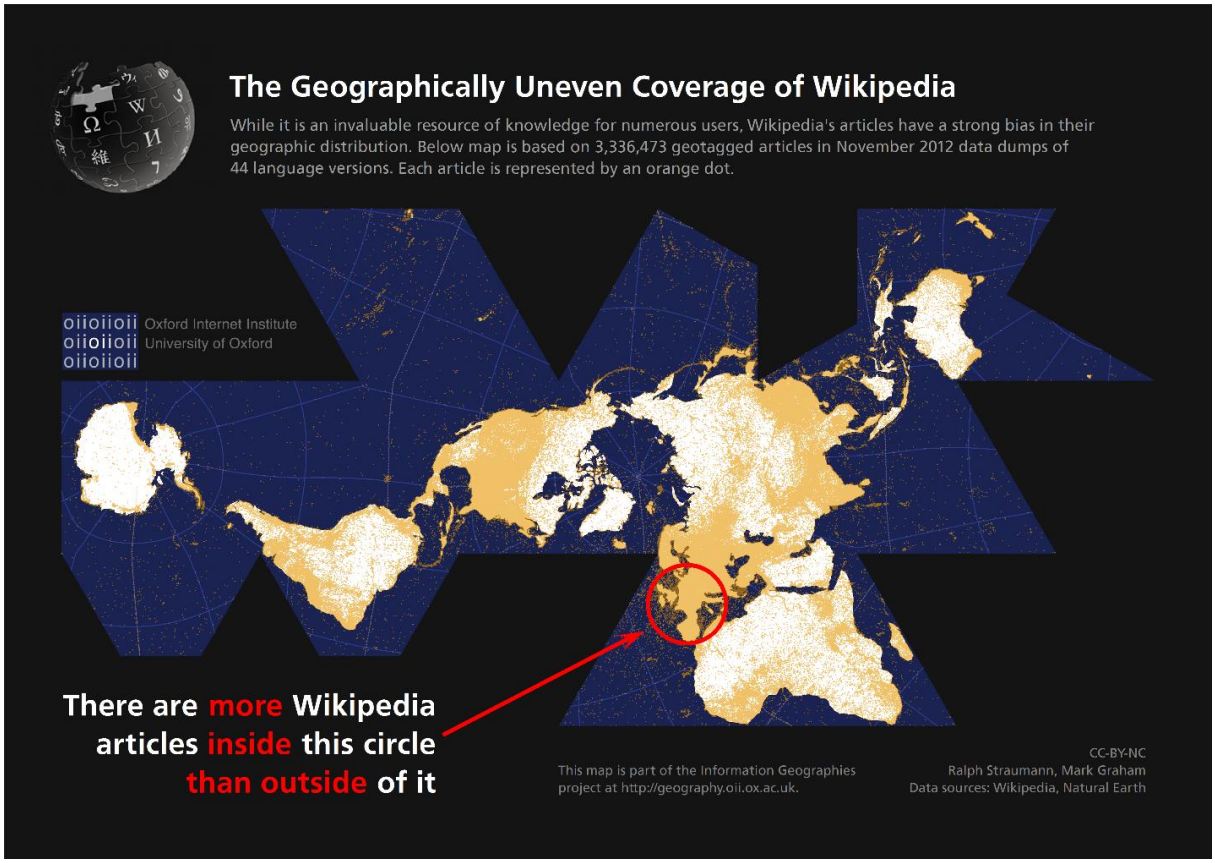


Figure 2.1 The uneven geographical coverage of Wikipedia (Graham & Hogan, 2014).

Wikimedia data dumps[20] are an important source for AI engines. As such, it is important to fill the gaps that might lead to a biased AI view of the world. Mozilla's Internet Health Report 2022[21] visualised Koch et al. (2021) research findings in an interactive map to show the origin of the datasets used in 26,535 research papers on projects using Machine Learning or AI performance benchmarking between 2015 and 2020. A large portion of this data originates

[20] https://dumps.wikimedia.org/
[21] https://2022.internethealthreport.org/facts/

from the internet, including Reddit[22] and Wikipedia, which have "encode biases potentially damaging to marginalised populations" (Bender et al., 2021, p. 610).

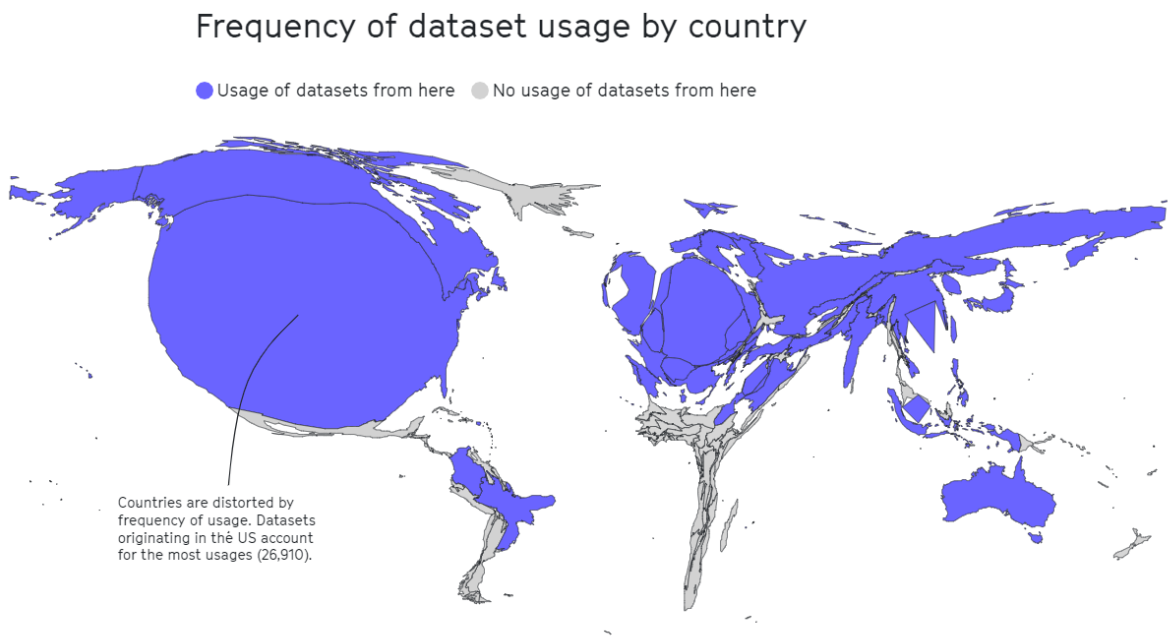


Figure 2.2 A map of machine learning usage of datasets made by Mozilla's Internet Health Report 2022 based on (Koch et al., 2021).

This uneven use of datasets indicates geographical disparity. A further challenge is the difference in quality, size, and user participation among the 300 different Wikipedia language versions, as quality standards differ enormously between Wikipedians[23] (Jemielniak & Wilamowski, 2017; *List of Wikipedias - Meta*, 2023). For example, 78% of Arabic Wikipedia is made up of stub (incomplete or too short)[24] articles (*Arabic Stub Articles*, 2023), compared with less than 40% of English Wikipedia (*English Stub Articles*, 2021). There are several culprits for this gap, ranging from digital literacy and limited access to the internet to social reasons such as conventions of collaborative practices prevalent in society (Pfeil et al., 2006; Morgan et al., 2012; Jemielniak & Wilamowski, 2017). Graham et al. (2015) identify three constraints that hinder local communities from geographical knowledge generation on Wikipedia: First, the absence of a local Wikimedia chapter or a user group to drive action and organise Edit-a-Thons[25] (an event where Wikipedia trainers, newcomers and experts community editors modify and enhance a particular topic or kind of articles on Wikipedia,

[22] https://www.reddit.com/
[23] Wikipedians is a term used to describe those who volunteer by editing Wikipedia articles. https://en.wikipedia.org/wiki/Wikipedia:Wikipedians
[24] https://en.wikipedia.org/wiki/Wikipedia:Stub
[25] https://en.wikipedia.org/wiki/Edit-a-thon

Section 4.4.1 includes further details). Second, the attractiveness of some topics about certain regions compared to others. Third, intense conflicts or disputes among editorial community members.

One of the largest programmes at the Wikimedia Foundation research team[26] created a roadmap to identify these gaps, recommend solutions, and invite other researchers from domains such as the HCI to collaborate[27] and understand biases and gaps in readership and contribution to building products that address them effectively (Redi et al., 2022).

Thus, addressing my research question and investing in developing collaborative solutions that bring change or complement the current models is crucial and needs to happen through the participation of diverse communities. The aim is to improve the onboarding and retention of editors who can contribute to filling the knowledge gaps.

2.1.3 Retaining Editors

Retaining newcomers is another pressing challenge for the Wikipedia community (Ciampaglia & Taraborelli, 2015). As English Wikipedia lost more than a third of its volunteer editor base between 2007 and 2013, the number of contributions declined (*Wikimedia Statistics*, 2022). Since 2015, the number of volunteer editors has plateaued (Erhart & Halfaker, 2015), with a slight increase during COVID-19 "mobility restrictions" in 2020 (Ruprechter et al., 2021) (see Figure 2.3).

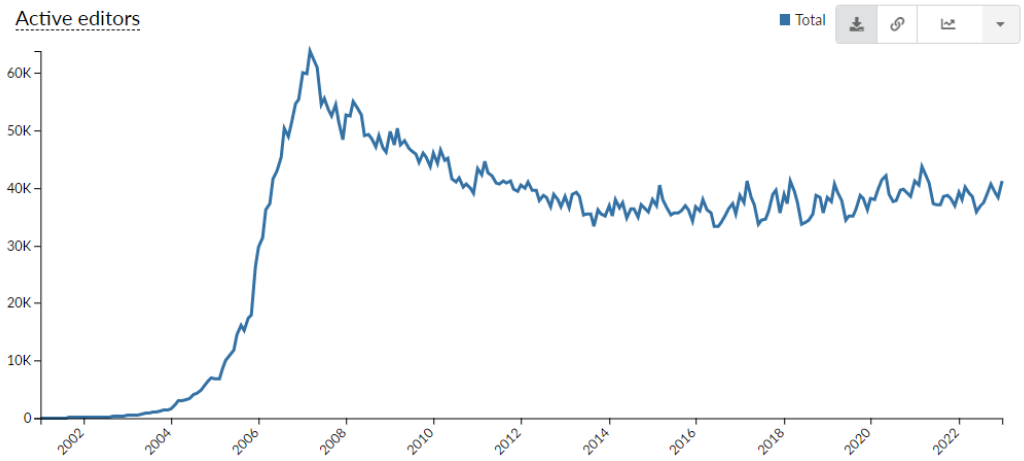


Figure 2.3 Active editors[28] in English Wikipedia.

[26] https://research.wikimedia.org/team.html
[27] https://web.archive.org/https://www.mediawiki.org/wiki/Wikimedia_Research/Formal_collaborations
[28] A registered, non-bot user who makes at least 5 edits to Wikipedia in a given. https://meta.wikimedia.org/wiki/Research:Active_editor

Covering the issue of the decline of Wikipedia editors from 2007 to 2012, TeBlunthuis et al. (2018) and Halfaker et al. (2013) attribute their low retention rate partly to some of the quality control measures imposed by the Wikipedia community. This reduces the likelihood of newcomer retention and increases the influence of members with a long editing history on the community rules and norms (Reboot, 2017). The authors conclude their study with a call for tool designers and community members to put effort into understanding the social norms and activities which those tools either encourage or impose. This is investigated and discussed in Chapter 7.

Morgan & Halfaker (2018) from the Wikimedia research team discuss the importance of providing Wikipedia newcomers with a positive onboarding experience and user-friendly tools to retain their contributions. They also discuss the increase in the number of returning editors after introducing the newcomers to the community norms in a friendly environment using a project called "the Teahouse". Morgan et al. (2013) describe the two primary motives for the Teahouse project as the "editor decline" caused by "the increasing rejection of new editors" and "a persistent gender gap". The authors emphasise the importance of engaging the community in a "bottom-up, community-driven" approach to finding solutions to these challenges. A redesign of the interface that does not respect the community history could disrupt the community or make the users feel manipulated and, in the worst-case scenario, could lead to "mass migration of long-time users from the site" (Gazan, 2011, p. 2855). This highlights the importance of involving the users affected by design in shaping it, which is the premise of my research project.

Similarly, Li and Farzan (2018) argue that supporting newcomers helps them overcome challenges and increases the probability of future contribution, compared to those who edit independently without community interaction. Analysing the edit history of Wikipedia articles on topics such as "Black Lives Matter" and "West African Ebola Outbreaks", it appears that editors motivated to join Wikipedia by an offline event produce lower-quality content and less sustained contributions than editors who have edited the same event pages two weeks later (Li & Farzan, 2018). This shows the complexity of retaining editors, understanding their motives, and designing onboarding experiences and collaboration opportunities tailored to their motives and needs.

The Wikimedia Foundation has made notable efforts to address these challenges. Such attempts manifest in their 2030 Strategy Recommendations (*Movement Strategy Recommendations*, 2022), which cover aspects such as improving the user experience to be inclusive. Additionally, Wikimedia aims to take an inclusive methodological approach to

improve its platforms and overcome impediments holding back its growth, diversification, and participation. One of their focus areas is adding communication features that support matching editors according to their fields of interest, motives, and areas of expertise to collaborate, interact with, and mentor one another (*Wikimedia User Experience Recommendations*, 2022).

I argue that efforts to meet these long-term challenges should examine the interplay of the different socio-technical components forming Wikipedia's ecosystem, starting with forming an understanding of how its technical infrastructure nurtures its community norms and how contemporary technological solutions can complement its ecosystem.

An overview of the collaborative Wikipedia editing infrastructure evolution discussed in the next section will help understand the obstacles hindering the adoption of real-time-collaborative writing and help explore ways for such a technology to contribute to solving those challenges.

2.2 Current Collaboration Model

Researchers consider Wikipedia a great example of a socio-technical system that orchestrates multiple components of technical infrastructure with the collective "wisdom of crowds", contributing to its content (Slattery, 2009; Niederer & van Dijck, 2010; Korsgaard et al., 2016; Tinati & Luczak-Roesch, 2017; Gheitasy, 2017). Contributions to Wikipedia are governed by detailed guidelines that have gradually evolved alongside the editing tools and features that Wikipedia provides.

This section provides a brief overview of the technical and governance infrastructure behind Wikipedia, focusing on how social elements have shaped the evolution of Wikipedia's technical infrastructure and vice versa. Then, a closer look is taken at the development of the Wikipedia community's forms of collaboration as a basis for the research question of how to responsibly introduce synchronous collaboration to Wikipedia while respecting Wikipedia's rich social structure and history. This would help answer how to introduce synchronous collaborative editing in a way that co-exists with asynchronous collaboration rather than as a replacement for current forms of collaborative editing.

2.2.1 Studying Wikipedia as a Socio-technical System

As a socio-technical system, Wikipedia has been studied from multiple perspectives, including quality, size, or other content-related topics (Kittur & Kraut, 2008; Arazy & Nov, 2010; Anderka et al., 2012), its community of contributors and their motivations (Kuznetsov, 2006; Balestra et al., 2017), and governance processes (Kriplean et al., 2008; Leskovec et al., 2010; Black et al., 2011; Müller-Birn et al., 2013). Views and research interests about Wikipedia are

changing over time as Wikipedia's social and technical environments evolve. This change is sometimes positive, such as the overall view on content quality. For instance, Murray (2019) encourages focusing health policymakers' and researchers' content creation efforts on Wikipedia rather than creating new online platforms since it has been evident that medical trainers at all levels use Wikipedia secretly for medical fact checks. On the other hand, feedback has sometimes been negative, as shown by opinions on topics such as gender and content bias, discussed in the previous section. Since these social and technical environments are inseparable, studying Wikipedia as a socio-technical system is vital and logical. My research project examines Wikipedia's technological and social infrastructure in supporting collaborative writing.

Collaborative writing groups often use many different writing strategies and document management methods across multiple technologies to achieve their goals (Baecker et al., 1993)**.** Baecker et al.'s findings show that the success of collaborative technology correlates with the ease of transition between collaboration modes and technologies. The socio-technical research approach to studying Wikipedia editing should acknowledge the fact that contributing to Wikipedia involves human interaction with various software and hardware. For example, during interviews conducted during my MSc dissertation project (Ardati, 2018) and in the sessions discussed in chapters 5 to 7, participants mentioned collaborating and communicating outside of Wikipedia, using tools and social platforms such as Google Docs, Facebook and Twitter. However, moving across various software and hardware for Wikipedia collaborators is further complicated by the fact that Wikipedia's UI is lagging behind these third-party tools (cf. *Wikipedia Gets a Fresh New Look*, 2023). The slow speed of innovation deserves study. My research argues that analysing the community's needs while considering the historical development of previously discussed challenges would help form a clearer picture to improve existing processes, tackle challenges, or change Wikipedia's collaboration model.

Connected Communities and Services

This section covers the complexities of such an analysis and shows the extent of Wikipedia's interconnectedness in the Internet ecosystem. It then focuses on Wikipedia's governance as another complication to Wikipedia collaboration, which dictates the socio-technical approach of this research.

Diving into the technical foundations of Wikipedia while paying attention to the social conventions and formations around the technologies within its community and other online communities is demanding interdisciplinary research. Researchers usually focus on a specific

element of Wikipedia independent of other factors at play in shaping Wikipedia, and in turn, missing the strong influence of the broader Internet ecosystem on its inner workings (Vincent et al., 2018).

An example of how Wikipedia is part of a wider Internet ecosystem is its Content Translation Tool,[29] which helped editors produce millions of translated articles in all languages by 70000 contributors (Ozurumba, 2021). This tool depends on the machine translation services of many providers, such as Apertium, OpusMT, LingoCloud, Matxin, and Google Translate. While its use is not uncontroversial, restrictions differ from one Wikipedia language community to another (*Content Translation Wiki Page*, 2022).

Using Google Translate is one example of how Wikipedia has benefited from using the surrounding ecosystem. In turn, it opens up the discussion around the possibility of adopting ubiquitous synchronous collaborative writing tools. Another thing to draw on from this example is how different Wikipedia language versions embrace the translation feature differently. This begs the question investigated in this research: Could this incremental or partial adoption apply to features such as real-time collaboration?

Studying Collaboration Governance

The Müller-Birn et al. (2013) study of the social norms and rules that govern the Wikipedia community points out that research on Wikipedia's social-technical structure focuses mostly on understanding how the coordination and communication of the distributed network of the Wikipedia community produce this massive online free encyclopaedia. However, this approach "does not fully embrace the socio-technical nature of the project" (Müller-Birn et al., 2013, p. 1). This is because Wikipedia is not only the result of extensive human efforts and millions of volunteer hours, but also the tools that support, manage, and maintain its content (Niederer & van Dijck, 2010). These latter experts are calling on researchers to investigate the socio-technical nature of Wikipedia and the close connection between the technological elements of Wikipedia and its community.

Therefore, a socio-technical approach is used throughout this research to examine the evolution of Wikipedia's social norms and its technical infrastructure for collaborative editing. This includes discussing the correlation between Wikipedia's technical and governance infrastructure, as well as a review of its software structure and related software products used by editors on participation and collaborative culture patterns.

[29] https://web.archive.org/web/20221221175806/https://www.mediawiki.org/wiki/Content_translation

The following section discusses the current Wikipedia infrastructure and how the community has changed their collaboration method over time.

2.2.2 The Evolution of Collaborative Editing in the English Wikipedia

Wikipedia is known to be "one of the most heralded success stories of peer collaboration" (Kittur & Kraut, 2008, p. 1), but what is meant by "collaboration"? And what are the different elements constituting Wikipedia's collaborative writing process? Unlike in CSCW, more generally, in collaborative writing research, the terms "collaboration", "coordination", and "cooperation" are often used interchangeably. This will become important in one of the sessions described in Chapter 6, where I discuss the implications of participants having different interpretations.

Mendoza-Chapa et al. (2000) define collaboration as the work of a group of individuals on a product that mostly involves making individualistic decisions that should fit the group's goals. They suggest that effective collaboration at a human level requires a high level of coordination of the group's resources and tasks to achieve its goals. It also involves the right interplay of the application-level components, referring to the different elements that form the collaborative writing application (Mendoza-Chapa et al., 2000).

Coordination has two main types: explicit coordination, which is manifest in the direct communication or planning between the editors using communication mediums such as the article talk pages, user talk pages, or a WikiProject page and implicit coordination, which occurs among a group of editors working without explicit communication but working on the same article and seeing each other's edits; and implicit coordination, which depends on unspoken shared understanding in addition to the right writing style for developing the article (Rouse et al., 1992; Wittenbaum et al., 2002; Kittur & Kraut, 2008). Over time, Wikipedia's community has documented in their various language wikis an extensive protocol that guides and governs editing processes for accomplishing almost any encyclopaedic task. Kittur & Kraut (2008) note that editing Wikipedia articles involves both explicit and implicit coordination, and both coordination methods are significant for successfully harnessing the "wisdom of the crowd". However, one type of coordination can be more appropriate than the other one based on different factors and circumstances, such as the group size and the stage of the article. Rezgui and Crowston (2018; 2020) identified a third variation to the explicit and implicit types of coordination taking place on Wikipedia, called "stigmergic" coordination. They based their findings on other research, such as Heylighen's (2006) paper, which explains work in a

collaborative environment as "a stimulus ('stigma')" driving greater contribution by others. In the context of Wikipedia, editors are responding to other editors' contributions. In Section 2.5., stigmergic coordination is discussed further, as is how this variety of Wikipedia's coordination models dictates the need for exploring more collaborative writing models that fit their needs.

Explicitly coordinating the writing tasks in a group of editors is often critical to ensuring its high quality, especially in an article's early lifecycle stages (Kittur & Kraut, 2008). An in-depth discussion of how measures are used to assess the quality is not covered, as they are out of the scope of this PhD research. However, research does show that collaboration and coordination patterns directly impact the quality of an article (J. Liu & Ram, 2011). Rezgui and Crowston (2018) and Kittur and Kraut (2008) highlight the importance of explicit coordination for small groups in Wikipedia, as starting an article may require a high level of synchronisation in planning and editing. However, when group size increases, explicit coordination, which requires interaction among editors, is more time-consuming and complicated than implicit or stigmergic coordination among a group of people who depend on a shared understanding of the writing tasks (Kittur & Kraut, 2008).

Wikipedia research on explicit coordination is limited to coordination in Wikipedia Talk Pages (Kittur & Kraut, 2008; Viegas et al., 2007; Rezgui & Crowston, 2018; Crowston & Rezgui, 2020). In this research, I am particularly interested in group onboarding environments involving a higher level of coordination among participants, such as Wikipedia training, opening more possibilities to investigate the utilisation of coordination tools other than the Talk Pages—internal or external to Wikipedia's technology ecosystem. This requires a closer look at the factors governing coordination in Wikipedia and how it has evolved in order to see if it is possible to incorporate a new collaboration model alongside the traditional ones.

Bipat et al. (2018) discuss how the development of collaboration models in the English Wikipedia has been driven by changes in community behaviour and the way they use existing coordination tools. A case in point is Wikipedia's Talk Pages, which were initially used to explicitly coordinate and discuss the development of an article, and have transformed into a space where users with similar interests gather for general discussions around the article theme. Bipat et al. attribute this change to Wikipedia's increased bureaucratic governance, the prevalence of bots that impose rules, and even bots making changes that replace, in many cases, the need for a discussion among editors on how an article should be developed. Another reason is the increased adoption of new underlying technologies and system features that give more control to Wikipedia administrators to perform major changes. Bipat et al. (2018) argue that with these changes, many assumptions that have been made in the past of how collaboration

works, such as the one about the use of Talk Pages, do not apply today and that this "opens the door for the development and design of new collaboration models" (Bipat et al., 2018, p. 9). They also point out that further research will be needed on developing collaboration behaviour over time. Such a study from a socio-technical perspective could support the users, increase the quantity and quality of articles, and enhance the user experience.

This section discussed research on collaboration on Wikipedia, while the following section discusses Wikimedia's innovation strategies and the efficiency and dynamics of responding to Wikipedia users' needs. This helps to better understand how it has become more difficult for Wikipedia to keep up with the fast pace of user demands, such as the demands for real-time collaborative writing.

2.3 Real-time Collaboration Feature

Currently, Wikipedia only supports the asynchronous collaborative editing model. Existing research demonstrates that most contributors engage in editing articles individually (Jemielniak & Raburski, 2014). This model is praised for building Wikipedia as we know it. The question is, is an alternative one needed?

The Wikipedia collaboration model encourages a solitary type of work where individual changes are subject to being challenged and potentially reverted. This creates a low-trust environment (Jemielniak & Raburski, 2014, p. 97). Establishing a higher "sense of trust", according to Ciffolilli (2003), has the potential to help Wikipedia retain editors.

The low level of trust which describes this community, according to Jemielniak (2012) and Jemielniak & Raburski (2014), contradicts what Olson and Olson deem as the "key ingredient in collaboration readiness" (Olson & Olson, 2013, p. 46). The difference in the stance between these authors is the process, which is primarily asynchronous in Jemielniak and Raburski's and mostly synchronous in Olson & Olson's work. Trust in the process and procedures substitutes the trust in individuals in peer production organisations (Jemielniak & Przegalinska, 2020, p. 44). Although Wikipedia has more than 50 policies, with 150,000 words (Jemielniak & Przegalinska, 2020), newcomers need not learn them all. This, however, shows how difficult and time-consuming it could be to find one's place in the community (Reboot, 2017).

Bryant et al. (2005, p. 7) argue that moving toward "fuller participation is becoming aware of the community you are joining." They found that through the transformation activity to become a Wikipedian, editors' view of Wikipedia changes over time as they progress from

novice to expert. However, they also find that Wikipedia supports expert Wikipedians' work more than the novices, although the latter might have changed over the past two decades.

Different motives to go through the journey mean that participants will have different levels of tolerance for the time it takes to trust the process and understand that they are part of a larger community. Balestra et al. (2017) address a similar question about the effect of motives by researching the difference in early motives for the editing activity over time and shedding light on the importance of early experience in retaining contributors. My research builds a case for the need to review assumptions about the "one size fits all" and study the situations when it is reasonable to provide different collaboration models for individuals with different motives.

It is important to take into consideration that some may want to "dip in" alone and edit as they go along, and that many open-source projects are actually driven by solitary contributors. Others might prefer to engage with a community and contribute through sustained interactions with other people more directly.

The following sections move to the technical side of this PhD research question by critically analysing why real-time collaboration is unavailable on Wikipedia. It examines its technological infrastructure, as well as previous attempts to add this feature, which social-technical complications have hindered.

2.3.1 Wiki Software

Real-time collaborative editing is not supported in Wikipedia. The last attempt to introduce it was made in Wikimedia's OfficeWiki,[30] the Wiki used by Wikimedia staff, but it was put into "Stalled" Status in 2018 with the following comment by a Wikimedia staff member: "As far as I can understand, it cannot be just enabled [at the moment], some development work is still needed. Stalling and marking as low priority then." (*Wikimedia's Phabricator*, 2022).

To understand why contributors cannot write articles in real-time on Wikipedia and how Wikipedia's technical infrastructure influences its collaborative community behaviour and norms, one must look at Wiki technology, which forms the core of Wikipedia website technology. Wikipedia operates using a customised open-source Wiki software called MediaWiki,[31] which the Wikimedia Foundation (Wikimedia or WMF, for short) supports. Now, Wikimedia provides the online encyclopaedia Wikipedia and its sister projects, such as Wikidata,[32] Wikiversity,[33] and Wikibooks,[34] with its infrastructure. By definition, a Wiki,

[30] https://office.wikimedia.org/wiki/Main_Page
[31] https://www.mediawiki.org/wiki/MediaWiki
[32] https://www.wikidata.org/wiki/Wikidata:Main_Page
[33] https://en.wikiversity.org/wiki/Wikiversity:Main_Page
[34] https://en.wikibooks.org/wiki/Main_Page

invented by Ward Cunninham, is a web-based software that permits anyone with access to alter its content using a browser (Ebersbach et al., 2008). Wiki technology is, in effect, built to serve as a quick way for anyone to collaborate on creating web content. However, multiple contributors submitting their Wiki edits simultaneously would trigger editing conflicts, the resolution of which can be time-consuming and frustrating (Rubart, 2017; Pace et al., 2018).

The next section explores how researchers and practitioners have developed the basic idea of the Wiki to serve both asynchronous and synchronous collaborative group editing. It also covers how these attempts influence the development of extensions with synchronous features to MediaWiki,[35] the Wiki software that serves Wikipedia.

2.3.2 Synchronous Collaborative Editing

Online collaborative efforts can be divided into four categories, according to time and space (see Figure 2.4): whether the collaboration is face-to-face or remote and whether it is synchronous or asynchronous (Ellis et al., 1991; Grant, 2001; Larusson & Alterman, 2009).

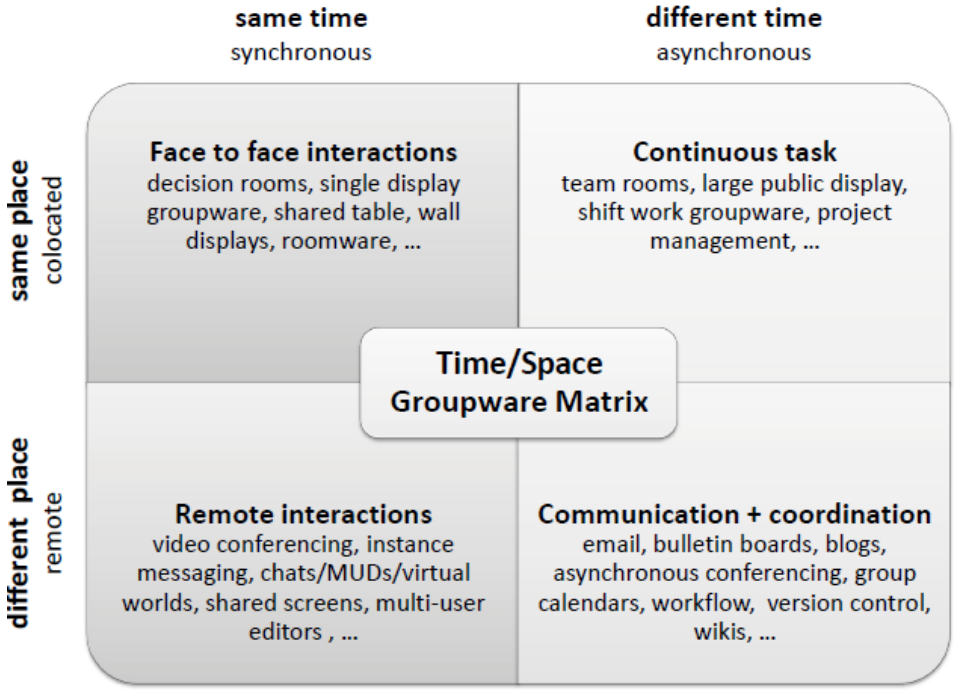


Figure 2.4 Time-space matrix for collaborative supported cooperative work adapted from (Johansen, 1988).

The boundaries are not necessarily clear cut. Google Docs recently introduced an offline editing feature allowing users with no Internet connection to edit a version of the document on their local computers ahead of merging it with the main document when the Internet connection

[35] https://www.mediawiki.org/wiki/Download

is available. However, asynchronous tools such as Wikis do not necessarily support synchronous features.

With today's tools, such as Google Docs and Etherpad, users may edit the same text simultaneously. A wide range of audiences use such tools to edit text together in real-time, adding greater simplicity to collaboration and allowing users to avoid the asynchronous model's unnecessary conflicts when multiple users submit their edits simultaneously (Pace et al., 2018).

2.3.3 Asynchronous and Synchronous Editing in a Collaborative Environment

The prevalence of commercially available collaboration tools has allowed users to collaborate on online writing tasks using the standard asynchronous mode and, currently, the synchronous mode, opening up new possibilities for working together (Olson et al., 2017). The authors find contradicting facts in their case to previous literature (Noël & Robert, 2003, 2004), which assumes that collaborators tend to fall back to the asynchronous collaboration model when offered both modes. Instead, Olson et al. find "strong evidence for people writing simultaneously, as well as hand-offs" and that simultaneous editing is a distinct plus to the collaborators' experience, as "some of the teams only worked simultaneously" (Olson et al., 2017).

My PhD research explores the assumptions made about Wikipedia's collaboration and the feasibility of merging synchronous and asynchronous features in a Wiki editing environment to help collaborators benefit from their volunteer experience.

Many educators have used a stand-alone Wiki as a collaborative classroom educational tool for assignments that develop students' technical and academic abilities (Rodero, 2017; Hudson, 2018). Other educators have experimented with extending their stand-alone Wiki Software features, using the extensive library of previously developed Wiki extensions, or designing their own. For example, Rodero (2017) studied the incorporation of Wiki functionality inside the Moodle learning platform and found that "[a] wiki can promote effective collaborative learning, confidence in self-training and peer evaluation, facilitating rapid feedback, intuitive navigation and task performance". Other educators use Wikipedia in the classroom to benefit students and society (Lampe et al., 2012; Konieczny, 2016). However, those who use Wikipedia as an educational tool cannot extend Wikipedia's functionality to accommodate their case-by-case needs because it is managed by Wikimedia and the Wikipedia community.

Researchers, such as Rubart (2017), have proposed design considerations to incorporate synchronous collaborative editing technologies into Wikis to enable multiple users to edit text simultaneously, thereby increasing the editing rate. The researcher's motive was to improve the Wiki text's quality, which positively correlates with the increase in the iterative collaborative process of editing and revising (Wöhner & Peters, 2009). Another example of a proposed collaboration model redesign is a prototype system developed by Biuk-Aghai & Lei (2010), who integrated an instant messaging feature that enables collaborators to communicate in real-time while viewing their Wiki page. They suggest that collaboration is a mix of synchronous and asynchronous activities; therefore, Wikis offering only asynchronous support should be complemented through synchronous tools that provide real-time feedback (Biuk-Aghai & Lei, 2010; Exter et al., 2012).

2.3.4 Wikipedia's Attempts to Include Real-time Collaborative Editing

The following section discusses attempts to extend Wikipedia's functionalities through synchronised collaborative editing.

In Wikipedia, the closest form to real-time collaboration is a group of editors explicitly coordinating article development through the Talk Pages and WikiProject,[36] which, according to Jemielniak and Raburski (2014), "rarely" take place and are "ineffective, due to the contingent and transient nature of its members". Further, they argue that what develops Wikipedia articles is the high number of editors and traffic on Wikipedia rather than its editors' cooperation. Kimmons (2011) goes further with their findings to suggest that the Wikipedia writing model does not reflect wider collaboration and heterogeneity in content building and macro-structural editing, "leading to the conclusion that most articles in Wikipedia are not reflective of the collaborative efforts of the community but, rather, represent the work of relatively few contributors." (Kimmons, 2011, p. 1). This can be seen in implicit coordination, where a few editors lead in setting the article direction, scope, and framework for other less involved editors to contribute (Kittur & Kraut, 2008).

However, the dependence of implicit coordination on shared understanding in a group would impose difficulties for less experienced editors to navigate through what the lead editors assume novices know about Wikipedia editing and article development (Rezgui & Crowston, 2018). The authors note that the editors' disparate expertise and knowledge levels and the

[36] https://en.wikipedia.org/wiki/Wikipedia:WikiProject

absence of face-to-face interaction aggravate this problem. Nevertheless, most communities do not meet to collaborate face to face, except for the Wikipedia training sessions run by Wikimedia and its community. As will be seen in future chapters, these training sessions may involve explicit real-time coordination while editing asynchronously and avoiding real-time editing.

According to the Wikimedia Foundation and the Wikipedia community, real-time collaborative editing is missing for a good reason. The problem lies in the social aspect of adding such a feature to the Wikipedia stack of technologies, as well as in the community rules and licencing complications of introducing multiple authorship. Nonetheless, many Wikipedia groups use external real-time editing tools such as Etherpad to organise their collaborative work (D'Angelo et al., 2018). This sparked a discussion between the Wikimedia community and the Wikimedia Foundation developers in 2011 about the possibility of having a real-time editor inside Wikipedia rather than using external real-time collaborative writing tools (*Real-Time Collaboration MediaWiki Page*, 2022). Furthermore, the Wikimedia Foundation has made multiple serious efforts to bring synchronous and asynchronous technologies into co-existence. In 2012, Mark Holmquist, a Software Engineer at the Wikimedia Foundation, introduced a tool built on Etherpad called EtherEditor[37] for real-time Wiki text editing, which extends the MediaWiki tool to allow a group of editors to collaborate on editing a wiki page. At the Wikimania conference in 2014, a group of Wikimedia staff attempted to extend the MediaWiki project's collaboration capabilities using an open-source JavaScript library called TogetherJS.[38] In their last attempt to answer the calls for adding a real-time collaboration feature, the Wikimedia Foundation introduced a beta version tool called CollabPad,[39] which was later added to MediaWiki's Visual Editor.[40] However, this tool has issues that limit the editor's capabilities, such as missing undo/redo and citation features.

The tools discussed can only be added as extensions to the Wiki software MediaWiki and cannot be used directly by Wikipedia users. In fact, most new features that significantly affect the user experience must be tested in the Wikimedia Foundation laboratories before introducing them as beta features for users to optionally add to their user accounts. However, Wikipedia has more than 300 different language versions. As a result, some Wikipedia communities may refuse to add certain features to their version of Wikipedia or restrict their usage should they believe they negatively affect their online encyclopaedia. For example, the content translation

[37] https://www.mediawiki.org/w/index.php?oldid=2768994
[38] https://www.mediawiki.org/w/index.php?oldid=3825797
[39] https://web.archive.org/web/20221225025202/https://visualeditor-test.wmflabs.org/wiki/Special:CollabPad
[40] https://web.archive.org/web/20221031034829/https://www.mediawiki.org/wiki/VisualEditor/Real-time_collaboration

tool has been in beta release since its introduction in 2014 (*Content Translation Main Page - MediaWiki*, 2023). While only editors with at least a 30-day-old account and 500 edits are allowed to use the tool in the English version, new users are permitted to utilise it in some other language versions, such as Arabic. This further complicates the possibility of rolling out new major features with socio-technical implications for the Wikipedia community.

The Wikimedia Foundation's IT team has concluded that collaborative writing technology is much less of a problem than the social aspect that must be addressed when considering a shift in support of collaboration on Wikipedia. For instance, attributing author contributions is more challenging in synchronous collaborative writing, where edits are less discrete and more overlapping. In addition, users' contribution records should appear in ways that align with the community norms of attributing editors. Other issues include maintaining quality standards and vandalism, which is challenging to deal with, especially when it occurs in real-time or when hundreds of editors decide to edit trending articles (*Waiting for Real-Time Collaboration*, 2017).

As the efforts to develop synchronous collaborative editing tools for Wikipedia have come to a halt (*Real-Time Collaboration MediaWiki Page*, 2022), the risk exists that synchronous collaborative editing becomes for Wikipedia what the 'edit button' is for Twitter, a feature much in demand but which never materialised, because of prior commitments and design choices. Yet, the ramifications of the changes should be studied in depth. Unlike recent Twitter changes that have worked against the stakeholders (Toniolo et al., 2023) and pushed some communities to migrate to platforms such as Mastodon (Kupferschmidt, 2022), my research is motivated by the need to include stakeholders in a participatory process and bring marginalised communities to join in building Wikipedia for "a world in which every single human being can freely share in the sum of all knowledge." (*Wikimedia Vision*, 2022).

It might, at the end of the day, be challenging to reconcile the idea of synchronous collaborative writing with Wikipedia's fundamental methods. However, as discussed later, having such a feature as a complementary method used in specific contexts can address many challenges and open up the discussion collaboration model on Wikipedia.

2.4 Chapter Summary

Previous sections in this chapter have shown the emerging need for a new Wikipedia contribution method to address challenges, such as the evolving coordination and collaboration patterns and community demands discussed earlier. However, that might be complicated due to the diverse, interdependent factors involved in Wikipedia article writing.

Collaboration that produces high-quality articles demands the effective coordination of writers as well as the presence of several interdependent elements such as resource collection, writing style, content, and article structure. By analysing a diverse set of English Wikipedia articles, Crowston and Rezgui (2020) found that most of their sample article edits lacked explicit task coordination but showed a high level of stigmergic coordination, as evident in the follow-up edits, based on earlier contributions to the shared work. Moreover, the authors found a positive correlation between stigmergic coordination and an article's quality in minor edits (fixing typos and sentences) as well as that explicit coordination is associated with higher-quality major edits (adding paragraphs and articles). Writing Wikipedia articles should accommodate these different needs for minor and major contributions, as the writing process tasks in each coordination style vary. It should be noted that Wikipedia's current technological stack of features supports this type of linear stigmergic coordination with its change-tracking features that notify editors of changes, resulting in more changes (Crowston & Rezgui, 2020).

Additionally, the Wikipedia community norms and needs have evolved yet are unmatched by Wikipedia's current technology. Research, in effect, shows that the role of Talk Pages, which is the closest thing to synchronous coordination, has diminished, gradually losing their function of facilitating discussions about coordination tasks, such as content writing and structure, in favour of discussion around the article topic (Bipat et al., 2018). Previous attempts to introduce a more collaborative model to Wikipedia have come to a halt due to the difficulty of information transfer between Wikimedia and the Wikipedia community (*Wikipedia Flow Project Talk Page*, 2016; *Real-Time Collaboration MediaWiki Page*, 2022). However, the fact that they have occurred repeatedly over time suggests that there are at least arguments for introducing synchronous collaboration in certain circumstances.

Incorporating synchronous collaboration in Wikipedia is possible from a technical point of view and might contribute to solving some of its challenges. However, it raises many questions from a social perspective, such as the fact that the Wikimedia community's governance and norms have developed around technological features that mostly support liquid collaboration and stigmergic and implicit coordination rather than synchronous collaborative writing or explicit coordination of tasks. Further, introducing synchronous editing begs many social questions, such as who deals with real-time vandalism and harassment, who runs the session, how it starts, how it is structured, how it ends, and who the text should be attributed to. Thus, adding such a feature without addressing these underlying issues might face rejections or restrictions from the already overstrained Wikipedia community, as has been experienced with restrictions imposed on the content translation tool.

This review on Wikipedia's nature as a socio-technical system and small group dynamics of coordination and collaboration using Wikis or Wikipedia articles lay the foundation to understand the potential application and benefit of real-time writing technology in contexts such as Wikipedia training, Edit-a-Thons and WikiProject, which can address some of the aforementioned social challenges. It also sets the direction for Chapters 4 to 8, which explore accomplishing this delicate task with further qualitative analysis of the community needs, as well as the mechanism of introducing this technology in Wikipedia's editing events, as illustrated through co-design prototypes in a unique innovation process that can also be used in producing solutions for similar online communities.

Chapter 3 discusses the methodology used in this research, leading to an observational study in Chapter 4 that recommends several techniques to facilitate the knowledge transfer among the sessions participants covered in Chapters 5 to 7.

Chapter 3 Methodology

Chapter 2 has demonstrated the need for further research to understand different models and opportunities for collaboration in Wikipedia. It has reflected the need to investigate the uptake, adoption, and integration of new technologies and group dynamics to onboard new editors in Wikipedia communities worldwide.

My research has two primary aims: firstly, to co-design support for real-time collaboration in Wikipedia training; secondly, to develop a co-design approach that can be used in distributed diverse communities similar to that of Wikipedia's. The work, specifically on Wikipedia, serves to validate the more general approach. The dual aim gives rise to the need to attend to the specific circumstances of the Wikipedia community while ensuring that the co-design approach can be transferred to other contexts.

Choosing a suitable methodology is key to studying an online community with diverse needs and evolving social dynamics, such as Wikipedia's. This diversity is demonstrated by the fact that some users make sure to attribute their contribution to their real names, while other editors use made-up usernames or even edit without logging in, resulting in many contributions associated with only IP addresses. Similarly, the devices used to access Wikipedia differ widely, with mobile access becoming more prominent (see Figure 3.1), as does the quality of network connectivity that people have available (see Figure 3.2). This is not to speak of the wide variety of cultural backgrounds and the growing belief that "there's no one size fits all for everyone" when it comes to designing for the Wikipedia community (Ho, 2021).

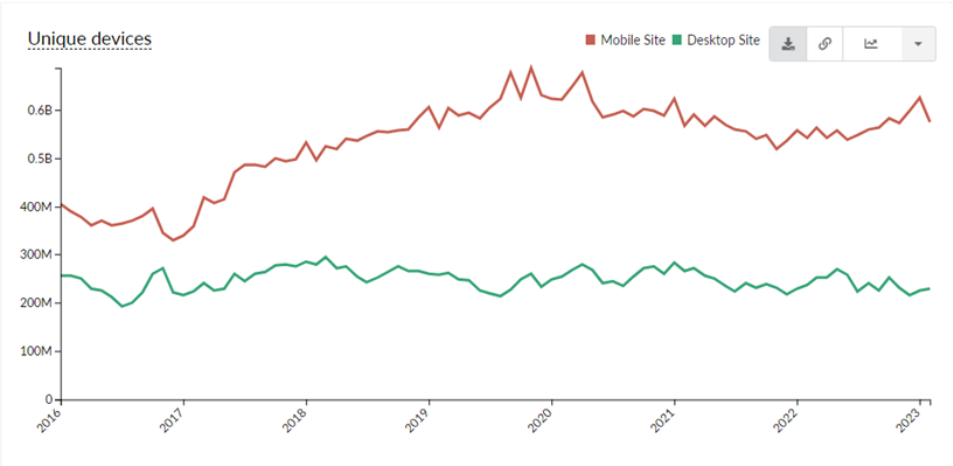


Figure 3.1 Number of unique devices accessed Wikipedia from a desktop or mobile device between 2016 and 2023. Taken from https://stats.wikimedia.org/

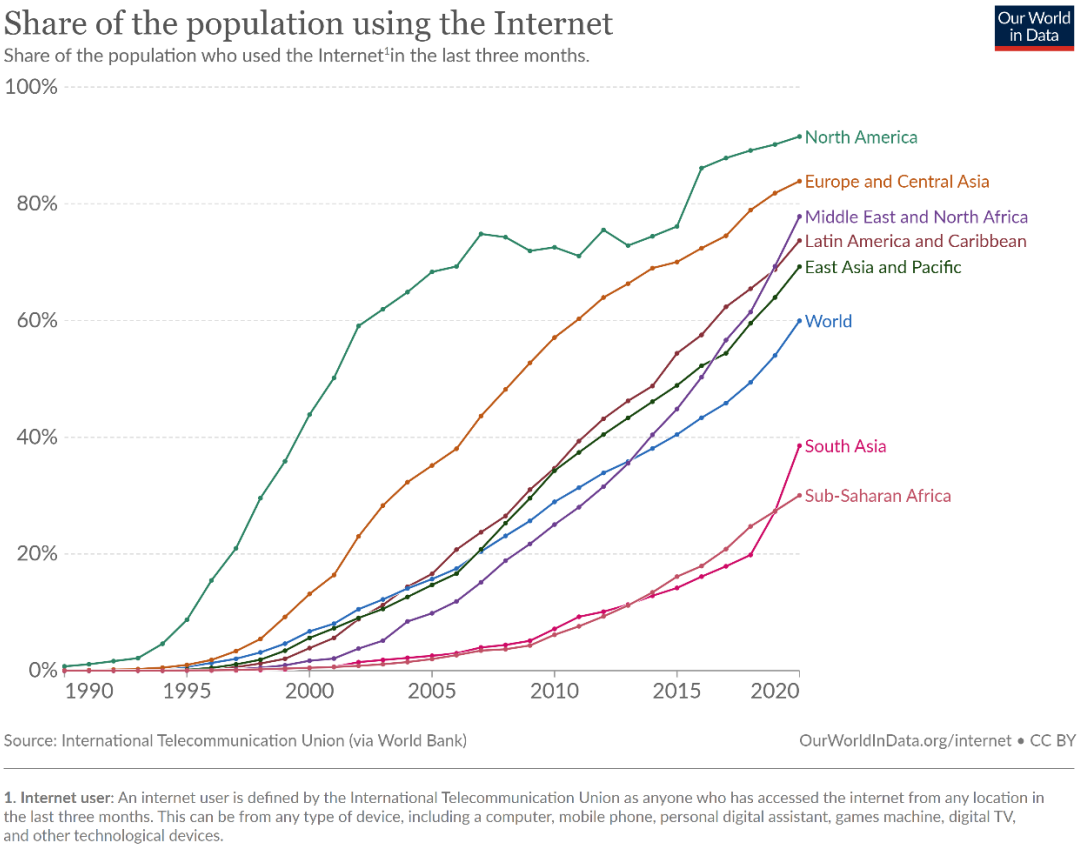


Figure 3.2 Share of population using the Internet between 1990 and 2020. Database Source: International Telecommunication Union (ITU) and visualised by OurWorldInData.org.

This leads to an emphasis on inclusion to ensure that different voices are being heard. On the one hand, depth in interaction with community members is required to ensure that, as yet, unknown issues are uncovered. At the same time, such depth can only be achieved with a relatively small group. Breadth is required to sample the wider community to guarantee inclusiveness and representation of a wider range of voices to weigh in on design decisions. These concerns shape the methodological choices with regard to the co-design strategy.

The aim is to contribute to the design community by producing a framework for a distributed co-design that can be used even by small teams with limited resources and increase the ability to transfer to settings that involve a similarly diverse, distributed community.

Sections 3.1 and 3.2 of this chapter provide a macro-level view of this PhD project's underpinning methodological stance and cover how events in the research context shape the approach and research focus through time to accommodate global changes largely brought about by the COVID-19 pandemic. While the concrete methods used are developed step-by-step, the overall methodological stance serves to ensure that they are cohesive. Section 3.3 shows the reasons behind selecting a mix of qualitative methods chosen at each point in time to answer the research question at hand. It then explores similar contexts and focuses on the

research work used to understand culture, design, and collaboration processes. Finally, Section 3.4 discusses my research co-design methodology.

3.1 Research Context and Scope

This section revisits the aim of this PhD and covers the development of opportunities in the research environment, reviews relevant methodologies, and establishes the foundation of using Ethnography and PD in building a new Ethnographically-informed Distributed Participatory Design (EDPD) framework. Appendix B covers the pivot from face-to-face to online due to COVID-19 as well as getting acquainted with the Wikimedia community and the training environment. It covers events and collocated training sessions that have given me a better understanding of the community's need for innovative collaboration models that meet the new challenges, influencing the research question. It includes a story of how participating in a training session to be an accredited Wikipedia trainer has helped define possible research collaborators and candidates for joining the design process. The new research direction for studying the community online also opened up the possibility of having much wider participation than just those near (or are able to commute to) the university facilities. This also shaped the path for reviewing the literature about design approaches with similar contexts, as discussed later in Section 3.3, seeking recommendations and shaping this research project methodology.

3.2 Methodological Stance

This section discusses the reasons why this research uses a qualitative methodological approach to better understand the absence of real-time collaboration and defines a framework to co-design a socio-technical system that supports real-time collaborative online communities.

A research methodology should illustrate a researcher's systematic approach to identifying and utilising a set of overlapping research methods and activities to address a specific research problem (Kothari, 2004). In addition, Pickard (2013c) discusses the importance of understanding the researchers' conceptual commitments and analytical stance to help guide the methodology while keeping the research aims firmly in view. The following sections cover the background and the rationale for choosing a qualitative, stepwise, and iterative approach.

Qualitative research is used "to answer questions about experience, meaning and perspective, most often from the standpoint of the participant [when] data are usually not amenable to counting or measuring" (Hammarberg et al., 2016, p. 499) or when not enough is known about what might need to be quantified. In the case of this particular research, there had

been no prior work on how Wikipedia training sessions are practically accomplished and how collaboration between trainers and trainees happens. It is impossible to see the actions of trainers and new editors purely in the data traces they leave behind on Wikipedia. Many of the activities in training occur in systems outside of Wikipedia itself, such as in video conferencing tools or other systems for coordinating the work. This made it necessary to start the PD process with an initial observational study and develop the methodological approach based on its findings.

In qualitative research, the research design is often not completely predetermined. Rather, it emerges as the research unfolds, something rarely "admitted outside the social sciences" (Pickard, 2013c, p. 16). Pickard (2013c) quotes Lincoln & Guba, who suggest an emergent research design for the researcher who embarks on a journey where the researcher "does not know what he or she does not know" (Lincoln & Guba, 1985, p. 209). Therefore, it is not possible to clearly define all methods to be used in advance to reveal and explore the 'unknown', which would only be discovered during the course of the research. Therefore, researchers are discouraged from having a detailed research plan ahead of time and are encouraged to develop their research design as the study progresses through an iterative process (Pickard, 2013c). "The research design must therefore be 'played by ear'; it must unfold, cascade, roll, emerge" (Lincoln & Guba, 1985, p. 203).

Wikimedia staff members attempted to introduce a new model for Talk Pages collaboration but faced many challenges, leading to attempts that ensure a higher level of community involvement (*Talk Pages Consultation 2019 - MediaWiki*, 2019). This affected the methodological approach in my research, which emphasises the importance of inclusion in representing a wide range of community members who can contribute broad expertise. However, the start of this research project did face challenges beyond the concept of inclusive participation, which is more about the project's practicalities. For example, it was difficult to define the type of different activities and schedule them to bring together a broadly distributed community, and recruit, establish, and retain a team with a united goal within a community that varies in motivation. The number of factors that were uncertain at the outset has made it necessary to develop the method according to the needs of the unfolding PD process, giving rise to a process that can guide future research in similar communities.

Research papers from the PD literature exclude a considerable portion of relevant information about how the PD is particularly accomplished as researchers tend to "simplify and abstract out the complexity of their contexts and problem framing, [and covering with limited details] how design, user involvement, collaboration, and participation are achieved or

not" (Botero et al., 2020, p. 17). Important details on how to accomplish PD work and transform a group of research stakeholders into a team to co-design a socio-technical system in a complex online environment therefore had to be worked out during the course of the research, leading to the need to iterate on the selection of methods and techniques used. Chapters 5 to 7 contribute to filling this gap by discussing in detail the strategising and mundane work my research has taken in preparing novel activities that build on the outcomes and emergent needs from prior sessions. These activities helped create an environment that fosters shaping the social and technical aspects of how real-time collaboration could be introduced.

The methodological stance that the method needs to suit the emergent needs of the project rather than being pre-defined is appropriate here for a number of reasons. First, participatory research and co-design must remain open to reflect the needs of the participants, which unfold only as the research gets underway. Secondly, the aim of this project is as much methodological innovation as it is to deliver outputs for the Wikipedia community.

Nonetheless, this research is undertaken with guidance and principles. Just as quantitative research has well-established criteria used to assess the results based on internal and external validity factors, so too is the qualitative work guided by criteria of integrity and trustworthiness (Pickard, 2013c; cf. Hammarberg et al., 2016):

- **Credibility**: Qualitative research should demonstrate a persistent engagement with the research participants over a significant period of time, using more than one technique to study the participants and their actions. In participatory research, credibility can also be established through feedback from participants and from peer researchers.
- **Transferability**: The ability to transfer the result to be used in a similar context.
- **Dependability**: Ensuring that appropriate methods are chosen and applied adequately.
- **Confirmability**: Being able to track findings and subjective knowledge of a researcher back to the raw data to make sure it is not merely based on the researcher's assumptions, interests or direct interpretation of the world being observed.

These criteria guided the process that led to the delivery of my research contributions. The research process for qualitative, participatory research is discussed in detail by Pickard (2013c), taking the four trustworthiness criteria into consideration.

The design of qualitative research involves the concept of the human as an instrument (Lincoln & Guba, 1985) to describe the human and her/his abilities as society members to interpret social phenomena as the primary instrument for scientific investigation. It encourages viewing the researcher as a human instrument capable of choosing and performing suitable

data collection methods supported by the researcher's growing understanding of the social phenomena being studied and of the needs of the participants who are able to make sense of their actions, which qualifies them to join the designers in the innovation process, for example, designing ICT systems (Heiskanen et al., 2010).

Purposive sampling is utilised to ensure the inclusion of extreme and typical cases, engaging multiple viewpoints using appropriate data collection techniques (Pickard, 2013c). Pickard describes the inductive data analysis that follows the sampling as an essential stage that results in ideas and concepts which require further investigation and theory formation, sampling, and data collection phases. Each iteration of this process should produce preliminary findings, which are presented and discussed with the research participants. The analysis in this iterative process leads to identifying themes that form the basis of a theory for the solution to transform from the local context of the research to the global stage (Pickard, 2013c). This should be an iterative process that stops when the data becomes repetitive, leading toward constructing the research outcome, which is later verified with the study participants.

Section 3.4 covers how my co-design methodology adopts this iterative qualitative approach in conducting an intervention-oriented participatory design to address the research question. As discussed later in Chapters 5 to 7, a vital factor for my research is this flexibility in having iterative cycles in which I draw on my research collaboration and experience with the community in reviewing the emerging data from the research to assess the need for methods and analytical tools.

3.3 Methods

Based on the context and evolution of this research, the following two main points shape this PhD project's co-design methods:

1) It is essential to understand the Wikimedia innovation process and the Wikipedia training context. This has led to the use of an **ethnographic observational study,** covered in Chapter 4.
2) Long-term engagement of stakeholders affected by the solution in the design process is paramount. This has led to using **a distributed PD process,** covered in greater detail in chapters 5 to 7.

The sections below give a general background of the methods used in different parts of the project. Further discussions of details will then follow in the coming chapters.

3.3.1 Ethnography

Research that helps understand users "will remain at the very heart of the computer's development. It is the engine of change." (Grudin, 1990, p. 267). Lazar et al. describe how designing a solution might require a "deep immersion and participation in a specific research context to develop an understanding that would not be achievable with other, more limited research approaches" (Lazar et al., 2017, p. 143).

The ethnographic approach to investigating a context for design in the HCI and Information System (IS) domains has become prevalent (Lazar et al., 2010; Pickard, 2013b; Hyysalo et al., 2016). However, ethnographic observation alone is not a suitable tool for predicting or constructing the future, rather, it serves to understand the practices of a community as a precondition for a design that reflects the concerns and practices of the members of a setting or community. In addition to providing concrete insights into social factors that constrain or enable design, it provides "models for thinking about those settings" (Dourish, 2006). Hyysalo et al. (2016) note that special attention "should currently be given to the ethnographic observation of real-life settings, as well as digital observation methods. These types of observation have now become common practice in mediating between design and use" (Hyysalo et al., 2016, p. 12). Understanding real-life settings is the focus of Chapter 4, whereas chapters 5 to 7 cover profound discussions with the community, which is helped, in part, by this understanding.

When "investigating" users as part of the Human-Computer Interaction (HCI) study, Hyysalo et al. (2016) call for gaining a detailed understanding of their behaviour. Especially when a system departs from its envisioned way of use, redesigning interfaces and interactions with the system becomes important. This adaptation of the system to its new contextual use is discussed earlier in Bipat et al.'s (2018) study on the emerging collaboration models (see 2.2.2) as well as the importance of critically revisiting the onboarding experience in Wikipedia training since the context of use has shifted online. This demonstrates a need for a digital observation study that could feed into the sessions for co-designing new training process with collaboration models that fit the Wikipedians' emerging needs.

While designing new technology would inevitably change some practices, the community's values and concerns, rather than specific practices, limit the designs. Therefore, despite being a member of the Wikipedia community for a number of years already, the observational study helps specifically with understanding the Wikipedia trainer community norms to guide the sessions. This PhD research is concerned with "understanding systems

within the situated context of the work setting" (Randall et al., 1995, p. 325) to co-design new work practices that support existing needs instead of remodelling the work process for introducing new features. In other words, the aim is to facilitate a co-evolution of the training practices, norms and proposed IT system rather than replacing the existing practice to introduce real-time collaboration. An observational study was, therefore, needed to allow me to observe training sessions with which I had little prior experience. I chose a non-participatory mode of observation so as to be able to pay full attention to the details of how the training sessions were accomplished that might otherwise have been missed.

Thus, developing a deep appreciation and understanding of the context would help me lay out the iterative process of engaging different research stakeholders, resulting in an outcome that suits the community needs.

Observation

This section discusses the observation method to collect data for answering questions associated with Wikipedia training, which is covered later in Chapter 4.

Observation is a prominent qualitative tool, the "principle technique of ethnographic research" to understand communities and "gain a close and intimate familiarity with a given area of study" (Marshall, 1994, pp. 158, 380). It is, in effect, one of the leading ethnographic data-gathering methods, a reliable method to construct a deep understanding of current systems' usage context (Hughes et al., 1992b, 1992a; Shapiro, 1994; Salvador & Mateas, 1997; Clarke et al., 2003; Randall et al., 2007; Lazar et al., 2010; Crabtree et al., 2012b, 2012a). In HCI, observation as a research tool is often combined with the researcher's participation in the studied group to generate rich data about their activities, interaction norms, beliefs, assumptions, work environment, and goals. Dourish (2006) attributes the adoption of this category of methods to the rise of Computer-Supported Cooperative Work (CSCW) studies that emphasise the social aspect of group activities and methodologies used to understand their social behaviour. However, there is a continuum from full participation, in what is called "observer as participant", to passive observer. The appropriate level of involvement will, inevitably, depend both on the research questions and on what the setting observed permits. Sometimes full participation is impossible (think of studying air traffic control), at other times, the presence of a passive observer can be inappropriate. In this specific case, since I had already attended several training events and was generally familiar with the Wikipedia community, I chose passive observation, which was unproblematic in the setting of online Wikipedia training sessions. This allowed me to pay specific attention to how training was practically

accomplished, both from the side of the trainers and the trainees. I also had been trained as a trainer, joined events as a co-trainer, and designed and developed training sessions. These different activities helped me understand the community's challenges, frustrations, and aspirations. Passive observation of training events in my PhD research helped me take a step back from this community membership and avoid projecting my assumptions to study the Wikimedia innovation process and the Wikipedia training from the research point of view without being pulled into a discussion about problems or challenges that do not connect to the research question.

The observational study covered in Chapter 4 helps form a broader picture of the innovation process as well as Wikipedia training activities, challenges, and opportunities, feeding into the PD co-design sessions, which involve direct interactions with the participants.

3.3.2 Participatory Design

Understanding context must be complemented with a method that guarantees sustained and meaningful interaction with the research participants as the research process moves towards design. This section covers some of the history of user involvement and collaboration with users in the design process, focusing on IT system user relations with designers or IT specialists. Furthermore, it addresses the choice of research methods, and the principles of user engagement used to generate a sense of ownership over the PD process. More literature is to be covered in Chapters 5 to 7 about the design methods or the analytical frameworks relevant to the studies, emergent data, and decision-making process, leading to their use.

Since the computer has moved gradually out of the workplace, becoming first a household item and now increasingly ubiquitous, designing software to cater to growing needs and diverse usages has become a challenge that has attracted researchers from several domains. Gradually, those researchers have collaborated on creating interdisciplinary research methods that factor the human into shaping the social and technical aspects of products and services. Consequently, researchers have become interested in users as producers of ideas, the solutions they can devise, the systemic ways that have rendered them productive, and the way co-design has been facilitated. Hyysalo et al. (2016) review user-designer relations in the twentieth century, which shaped users' role in innovation and design. They discuss the shift in the late 20th century from analysing the users as *objects* to exploring their role as *subjects* of collaboration on innovation, leading to the emergence of such new design strategies as collaborative design and human-centred design (HCD).

During that period, PD emerged as a methodology from the trade unionists' needs to protect their work through direct engagement in design decision-making and implementation of companies' technological projects. It has evolved as an inclusive design approach since its use in the Scandinavian workers' rights movement in the 1960s (Greenbaum & Loi, 2012; Kensing & Greenbaum, 2013). This sparked the development of many collaborative design techniques, especially in the IT sector. Domains such as human-computer interaction (HCI) embraced the PD practices, contributing to its development with various user involvement methods in information system (IS) design. The usability of a software application or computer system has always been one of the main concerns of the HCI community. Participation of system end-users in the design process has been proven to contribute to a system's success (Allen et al., 1993). Danielsson et al. (2008) explain how this has led to the wide adoption of practices that involve the system end-user in the development phase for a more usable and efficient technological solution.

For my research, PD provides an opportunity to bring the community together to discuss training, opening a space for reviewing assumptions about the collaboration model and leading to incorporating real-time collaboration as an alternative writing model. However, the usefulness of such an opportunity to involve the community is contingent on establishing relations that foster collaboration and knowledge transfer between users and designers (Heiskanen et al., 2010). Engaging end-users, however, is not a straightforward task, as merely bringing users together with researchers or designers does not necessarily produce a more user-centric design for the system. It requires more than what Muller and Druin (2002) refer to as "just add users and stir." For researchers to design a successful system for all stakeholders, the need exists for methods to create an environment where the power of influence is shared among an acceptable proportion of a system's user-base community and researchers (Muller & Druin, 2002). A co-designed system should fit its users' socially organised work. To meet this requirement, Voss et al. (2009) outline basic principles for user-designer relation:

1) Work is socially organised and occurs in a certain situational environment; as a result, so does IT systems usage.
2) Design must be based on comprehension of the context.
3) Including users in the design process is often advantageous.
4) Users possess skills and resources that are important to the design process.
5) The introduction of IT systems alters the function of certain skills.
6) IT systems should support working practices and quality work, rather than just quantity.

7) The design process is political, which may stoke conflict.

These principles vary in relevance in different domains; however, they all emphasise the importance of understanding the context of users and the direct engagement of users in shaping the systems that will affect desirable sustainable solutions.

Researching a problem requires utilising a tailored set of methods that fit the researcher's needs to develop a solution based on concrete scientific findings (Kothari, 2004). Therefore, this research needs an inclusive methodological approach that positions the user needs at the heart of the work. PD has been evolving as an inclusive design (Hyysalo et al., 2016), empowering end-users by increasing their influence on system design and including their voices by working with them directly to understand their social and technological needs (Greenbaum & Loi, 2012).

Greenbaum & Halskov (1993) discuss how the PD research approach is renowned for its unique, pragmatic methods and for emphasising distributing the decision-making power among all those affected by the solution design. Furthermore, it fosters an environment to harness stakeholders' political motivation and empower end-users to express their thoughts and expectations of their future. The other important element of PD is its aim to draw on the end-user's practical knowledge about established practices and their views on the tools used to accomplish their tasks.

Many methods can be used in a PD study; that said, finding methods that suit a specific setting for a particular purpose and at a specific research stage can be challenging. Therefore, the Young and Well Cooperative Research Centre, with its partners, have developed a framework (see Figure 3.3) for guiding researchers on the use of PD in community research projects to bring about the "development of interventions that are effective, relevant and appealing." (Hagen et al., 2012, p. 1).

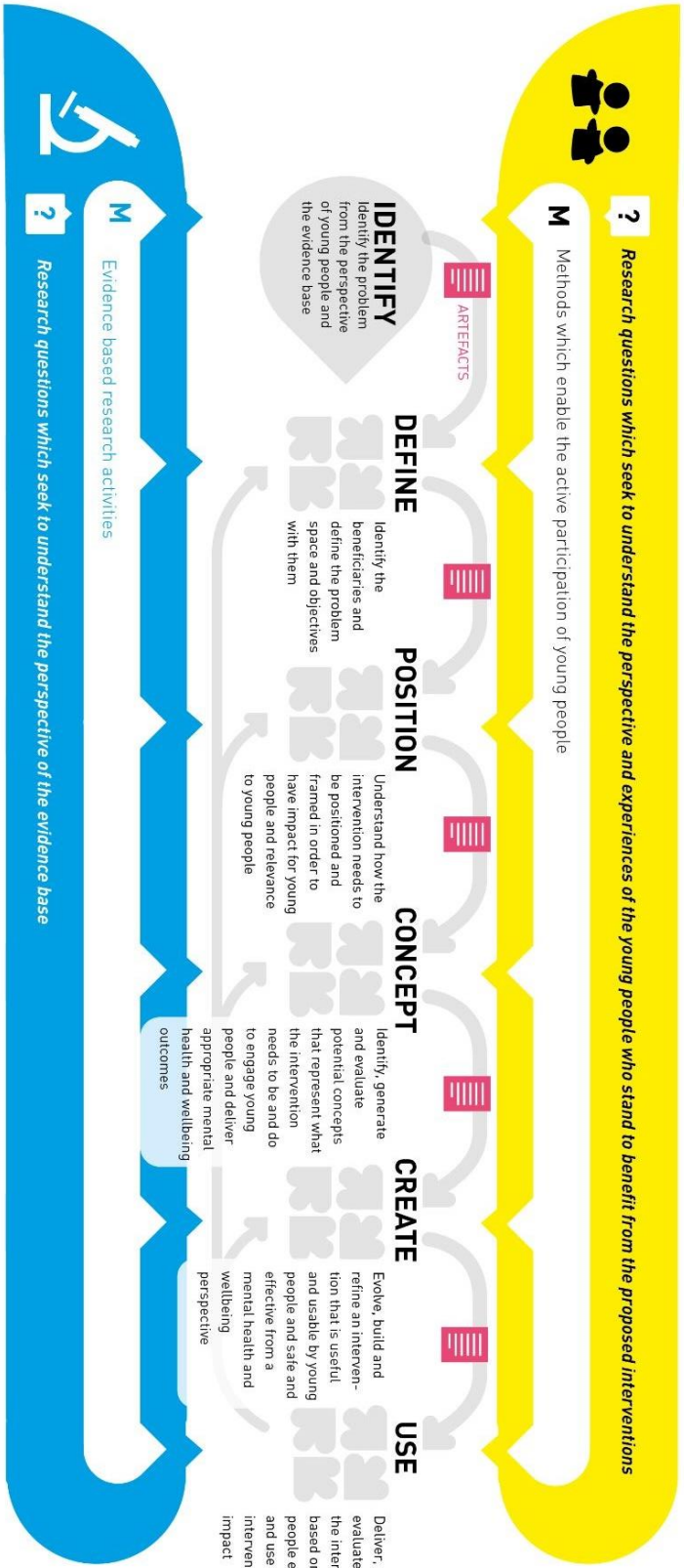


Figure 3.3 The components of Hagen et al.'s (2012) framework.

With Australian Government support, they run an Internet-based initiative to address the high suicide rates among people aged 14 to 25 through online intervention. The idea is to encourage researchers and institutions to build online platforms to support communities using a PD approach. Along with several case studies, they offer an emerging set of critical, conceptual, and practical tools that support the active participation of users in the design of platform systems and services (2012). These researchers discuss how the PD approach overlaps with the standard user-centric design (UCD) methods and goes further than evaluating "what works" and "what doesn't work" to actively involve the users as co-designers throughout the design process. Hagen et al. consider PD a viable strategy for understanding and incorporating young people's views, experiences, and creativity that such intervention seeks to acknowledge and help. Their framework for practising PD aims to ensure a thorough investigation of what the issues are, how they can be defined and how they manifest in the lives of people. It highlights the need to comply with the following principles:

- End-users should be active collaborators (co-designers) throughout the design process, from problem definition through problem resolution.
- End-users should participate as design partners, contributing ideas as well as comments and feedback on current design concepts.
- Suggested solutions are examined and reviewed continuously in terms of relevance, meaning, and engagement with the end users who are affected by them, as well as their potential for causing tensions that may negatively affect the design stakeholders.

Hagen et al. (2012) consider their main research contribution to be a framework for researchers considering adopting PD to engage with online participants to provide challenging mental health intervention. Instead of reinventing co-design concepts and practices, Hagen et al.'s (2012) framework provides a flexible, phasic basis from which my research benefits. This PhD research adopts Hagen et al.'s framework to build a new one catering to widely distributed, large online communities while keeping the resources needed to a minimum. Therefore, my framework differs from Hagen et al.'s in several key respects based on the needs of my research context. These differences are briefly discussed below and covered in further detail in Section 3.4 and Chapter 8.

While Hagen et al.'s framework provides generic guidance to researchers on PD, the proposed framework in this thesis offers significant further development in terms of the breadth and depth of exploring the context, scaling user participation up and down, and digging into the root of the socio-technical complexities using a variety of methods and tools adapted to

navigate online communities, their social norms and practices. The new framework significantly benefited from using a combination of ethnography and participatory design in the context of the Wikipedia community. It demonstrates in detail how to closely collaborate with online communities on an ethnographically-informed PD approach to evolve their norms around technology, mapping and addressing constraints in the existing socio-technical infrastructure.

Adding the users to the design process, however, is inadequate, especially in exploring new innovative solutions that could radically change community practices beyond merely remoulding current practices. The proposed framework assures active end-user involvement and an empowered position in an open-minded, inclusive solution exploration in the design process, thereby balancing the influence power among the design process stakeholders, which, in this research case, are the Wikipedia community, Wikimedia Foundation, and the designers. Furthermore, the framework divides the design process into several phases that explore different social and cultural elements of Wikipedia's social and technical settings. This includes initiating the process by discussing broad topics about the community practices and challenges that do not necessarily directly target the project aims. This is followed by gradually directing the discussion flow towards testing the relevance of the need for changes that bring new socio-technical solutions to the community. It also includes iterating through these stages, when necessary, to guarantee that all stakeholders are involved in the ideation and evaluation stages that facilitate knowledge sharing among participants.

Hagen et al. framework use cases:

Hagen et al.'s PD framework has been used in a number of contexts. In response to Hagen et al.'s (2012) call for employing PD in the youth mental health context, Wadley et al. (2013) acknowledge that, due to the high rate of technology use among young people, HCI methods that foster better usability and acceptability would contribute significantly to solving such a stressful social challenge.

In a similar move to Wikimedia UK's effort in moving Wikipedia training from face-to-face interaction online, the researchers started their project based on the premise that technology could be utilised to extend and complement the existing face-to-face therapy model with online therapy delivery. According to them, this would mitigate the heavy toll of mental disorders on the young people of the Australian community (Mathers et al., 2000), especially patients who have busy schedules or are geographically disadvantaged in accessing traditional face-to-face mental health support (Beattie et al., 2009). Despite the challenges of conducting research in the online mental health context, and after three years of employing PD methods in

collaboration with a research-focused clinic, Wadley et al.'s (2013, p. 1) illustrate that PD research resulted in "technology that is efficacious and acceptable to young people." They held several co-design workshops solely with patients at the beginning and, later, with clinicians separately to develop a comprehensive online clinic solution which replicated some of the face-to-face patient visit models used in the clinical treatment sessions. For example, they emulated a card-sorting practice to help patients identify their strengths, weaknesses, and social conditions that lead to a warning or an uncomfortable situation. In addition, due to the sensitive nature of the research area, the researchers faced difficulties in implementing the PD; for instance, some patients did not want to be identified by or to identify other patients and did not want to commit to regular travel to workshops and testing labs. Wadley et al. (2013, p. 8) hoped that their design approach would help other researchers with similar "context in which research participants and reported experiences can be difficult to access." This paper's approach of adapting methods designed for face-to-face implementation online has helped my co-design approach, culminating in an engaging process in a context where the studied community can be difficult to access.

This is quite common in researching the Wikipedia community and getting them to trust researchers' intentions (Graham & Hogan, 2014, p. 111). In my MSc dissertation (Ardati, 2018), the community was sceptical and hard to recruit, with many wanting to preserve their anonymity. This was a concern for me, as I approached the community again for research with a longer time span and regular engagement, potentially causing them "research fatigue" (Clark, 2008).

Another study that came in response to Hagen et al.'s (2012) call for projects that support youngsters with mental health online is the Ospina-Pinillos et al.'s (2019) study, which used participants in sketching interfaces to co-design an e-clinic application with multi-language support. This study involved a translation and cultural adaptation phase since the final product is directed to non-English speakers living in countries where English is a native language (Ospina-Pinillos et al., 2019). The resemblance to the diverse needs of Wikimedia's multilingual user base is quite evident here. Ospina-Pinillos et al. (2019) have contributed to my selection and elaboration of methods. For instance, similar to their approach, my research has adopted wireframes in prototyping and used NVivo software[41] to do the thematic analysis (TA) of the notes and the participants' artefacts from the workshop. They used the Valdez et al.'s (2012) Culturally-Informed Design Framework, which guides designers in developing a

[41] https://lumivero.com/products/nvivo/

user-centric IT health solution. The framework helps in the conceptualisation and categorisation of prototyping decisions into four main themes: the "technology platform", which represents the hardware to deliver the solution, the "functionality", which describes the features of this solution, the "content," which illustrates the textual data provided by the tool, and lastly, the "user interface" which describes how the functionalities and content would be presented to the end-user. As discussed in Section 4.4.3, this framework is partially used in my research for coding observational notes.

PIPWatch – a Distributed Participatory Design Case

Ospina-Pinillos et al. (2019) and Wadley et al. (2013) have developed online software solutions that target the healthcare sector but involve face-to-face interactions with participants in the PD process. However, my choice of methods is dictated by the restrictions imposed by the COVID-19 epidemic, combined with the difficulty of running inclusive PD research for the Wikipedia community, which, as mentioned earlier, has a complex online presence in terms of its geographical distribution and a need to preserve the anonymity of the community members who require it. Therefore, this research relies on methods and tools that compensate for this lack of face-to-face interaction between the designers and the end-users of the online solutions. And as discussed later in Chapter 8, it addresses the question of how and when to scale up the participation of the broader Wikipedia community in the design process, which is partially based on Clement et al. (2008) study as discussed below.

In 2005, out of interest in enhancing web browsing privacy, Clement et al. (2008) developed "PIPWatch," a Firefox extension to rate website privacy policies and practices. The researchers considered themselves end-users as well as developers. They identified the potential of PD research in developing "Web 2.0" applications. They wanted to provide an inclusive online environment that encourages active and continuous participation in the design process for users, regardless of their geographical location, while preserving their anonymity. Clement et al. reviewed several successful crowdsourced platforms in an effort to understand what drives people to voluntarily contribute their time and expertise in the form of textual content to websites such as Reddit. The aim was to use the same techniques in getting sustainable input from people in terms of designing and developing their privacy tools' socio-technical infrastructure. Their development process started with informally collecting requirements and suggestions about the tool's requirements, which they later organised using an online forum. Within the space of two years, they had multiple iterations of designing, prototyping and integrating the feedback of a growing active user base.

Clement et al. (2008) reviewed the development of popular "peer-production" platforms, such as Wikipedia, and found a lack of user participation in the services designing stage. They noted that most design changes in these platforms happened in response to analysing their user behaviour rather than direct involvement in the design process. Finally, they provided lessons learned from their distributed PD study, highlighting the need to organise the participants' contribution flow to keep the online discussion alive and drive more contributions. They also emphasised the importance of expanding the repertoire of PD methods and techniques to include features that celebrate enthusiastic contributors in the "Web 2.0" applications, characterised by an increased focus on user experience, interactivity, and collaboration. This demonstrated a need to review Wikipedia's latest approach to user participation in the innovation process, which is covered in Chapter 4. Clement et al.'s (2008) idea of having a public forum for discussing the design has contributed to the scale up in the Social Voting phase in my research, where the Wikipedia community is consulted on the ideas that have emerged from the sessions.

The aforementioned studies and the emergent need from my research produce a distributed co-design framework that can be used in distributed diverse communities, similar to Wikipedia's, which is discussed further in Chapter 8.

Distributed PD for Wikipedia Community

This section explores the relevance of using a distributed PD approach for creating an online co-design environment for the research participants.

Greenbaum & Loi (2012) explain how PD establishes a mutual learning environment that enhances the understanding of different system stakeholders, forming a common ground to build a workable solution for a broad spectrum of stakeholders. Muller & Druin (2002) define PD methods as a hybrid practice that takes place in an "in-between" region that shares the attributes of both the user and the technology developer's field of knowledge. The PD approach intersects, in turn, with HCI's calls for a "mutual or reciprocal learning in hybrid spaces" (Muller & Druin, 2002, p. 11). However, creating this space for a widely distributed community is difficult since most crowdsourced projects are set to serve a community that comes from a broad spectrum of cultural, technical, and social backgrounds. For example, Wikipedia is famous for its ethos that "Anyone with internet access can write and make changes to Wikipedia articles" (*Wikipedia - About*, 2022). Despite the difficulty, it is crucial to create this inclusive design space. Lack of inclusion in the early design phase of Wikipedia's socio-technical infrastructure, as discussed in Chapter 2, led to receiving contributions from a specific

portion of the global community. For example, the English Wikipedia average user is characterised as a technically savvy white man between 15 and 49 years old with a white-collar job, formally educated, and from a developed nation in the northern hemisphere (*Systemic Bias - Wikipedia*, 2022). According to (Ford & Wajcman, 2017), this bias has resulted in a considerable gender gap in Wikipedia editors' community and content. In addition, Wikipedia's editing is criticised for being too complicated for those with no technical background, resulting in calls to improve and simplify the user experience by the researcher community and its contributors (Ford & Wajcman, 2017; Gardner, 2011; Halfaker et al., 2013). Therefore, the design should involve people with diverse demographics to develop solutions that encourage a more varied contribution to Wikipedia.

PD was mainly developed to democratise engaging the workforce in large organisations to build more usable information systems for the workplace. That, however, does not cover the emerging needs of new organisational structures, such as online communities. To cope with these needs, PD has expanded its methods to deal with key stakeholders with different power levels in the community who are distributed across time zones (Obendorf & Janneck, 2009; Kautz et al., 2018). This distributed nature of work can be seen in open-source communities, collaborating on the continuous development of technologies such as Linux operating systems (Raymond, 1999).

Over time, the Wikipedia community has developed a unique model of engagement with one another and the Wiki technology that runs Wikipedia. The majority of community members are volunteers who have never met each other and are motivated by the desire to acquire and produce knowledge and improve Wikipedia to be recognised and respected as an accurate source of information. Most conventional PD methods are targeted toward users who can join design sessions in the same physical space (Clement et al., 2008). However, the research complexity is aggravated when researching online communities that are distributed geographically, and that potentially prefer to preserve their anonymity due to participating in controversial topics (Clement et al., 2008). To solve this problem, the distributed PD approach is intended to bring together a diverse range of stakeholders in distributed design teams. It provides convenient solutions and methods for teams collaborating on co-designing a solution through the Internet, web-based tools, and social media (Lukyanenko et al., 2016).

According to Panciera et al. (2009), Wikipedia's diverse editorial community have a unique complex participation pattern and life cycle. For instance, Panciera et al. (2009) show that the number of edits on the first two days of membership strongly correlates with the user's likelihood of becoming an active user in the future (Panciera et al., 2009). They also indicate

that the online community diversity, and its participation norms, which evolve, may influence the member experience, especially new users who want to join the community. A case in point is a study of the Python online community which shows that if a newcomer wants to become an active contributor, the community's "rites of passage and the canonical trajectory" are not fully documented on Python's Website (Ducheneaut, 2005, p. 352). Furthermore, Ducheneaut (2005) explains how very few newcomers to this community successfully evolve and develop an "insider" identity after undergoing a combination of enforced and implicit norms. In fact, one of the major challenges for Wikimedia and the Wikipedia community is this steep, bumpy learning curve to transition from being a Wikipedia reader, on the edges of its community, to the core ones (Bryant et al., 2005). This complexity and some contradicting aspects of the onboarding procedures are well documented in the research community, as discussed in Chapter 2, and are covered in depth and analysed in Chapter 5.

Collaborating on designing a solution for the Wikipedia community brings more challenges, such as coordinating tasks, communication, and conflict resolution, which most user-generated content platforms suffer. My research contributes to this area by proposing and applying a suitable distributed PD approach based on Hagen et al.'s (2012) framework and previous findings. The choice of methods is guided by the need to achieve inclusion in the PD exercise. On the one hand, it is essential to involve participants deeply in a co-design exercise, helping them use their expertise effectively. On the other hand, it is important to have broad participation to ensure a plurality of voices are heard and provide more comprehensive community support for the system being developed.

Thus, the study uses both synchronous methods that allow deep deliberation of design alternatives and asynchronous methods that require little time commitment and scale to the broader community.

3.4 Ethnographically-informed Distributed Participatory Design Methodology

Hagen et al. (2012) discuss how it is ideal to start the PD process at the inception stage, but also mention the possibility of starting PD intervention at any research project stage. In my co-design approach, I have added an observational study stage (see Chapter 4) before the inception stage and the start of the co-design sessions. The discussion of the research context and scope has already established the uniqueness of this project's research context in researching the online Wikipedia training community. This section discusses existing frameworks and how they have been tailored to this research environment, leading to the new proposed EDPD

framework. It covers how the process is altered to accommodate the web-based and distributed nature of my research.

A complex social environment, such as Wikipedia, that facilitates mass collaboration requires new innovative practices for finding software solutions. The rise of socio-technical peer production of information networks has given birth to new software engineering and design practices, such as the (Kazman & Chen, 2009) metropolis model that describes the relationship among three different roles in the development of crowdsourced systems. First is the kernel, which consists of the product owners, software architects, and policymakers, whereas the periphery consists of the developers and prosumers (those who consume and produce simultaneously). Lastly, the masses are defined as end users and consumers. Each category requires a set of tools and methods for efficient engagement in the product development process.

In my distributed PD approach, Chapters 5 and 6 discuss involving the kernel and periphery categories. According to Kautz et al. (2018), there are limited cases where PD has been used to facilitate cooperation between kernel and periphery on open-source tool development. Their research relied on email, telephone, Skype and Instant Messaging to communicate with core stakeholders and social media to get the input of the periphery. My research incorporates virtual collaboration boards as a medium for bridging the gap between the kernel (Wikimedia UK staff and veterans with administration power) and the periphery (experienced editors, trainers and newcomers). Furthermore, Chapter 7 covers involving all three roles, including the masses (newcomers, readers, and Wikipedia casual editors), in the design process.

In my work, I have had to research context where the engagement with participants had to be online, take into consideration varying levels of access to the internet, and must be responsive to participants' time constraints as well as accommodate different time zones. This meant that there was significant work involved in managing the practicalities of the PD collaboration. This involves both scheduling and providing summaries of what had been achieved in cases where new participants joined. This kind of work will exist in some form or another in many PD projects but is usually excluded from accounts in academic papers or reports such as Hagen et al.'s. The socio-technical complexity of the project means that tracking issues and accounting for design decisions is of particular significance. It has led to the usage of analytical tools in between sessions to analyse recordings of the sessions (see Thematic Analysis in Section 6.5) and to disentangle complex challenges (see Activity Theory

in Section 7.1). The results of these analyses needed to be turned into reports for the participants.

The result is an approach that remained flexible in choosing specific methods in response to issues encountered within a phasic approach adapted to researching online communities. The thesis provides a detailed account of the mundane work of PD and of the choices made. The overall methodological stance can guide other researchers through similar complex socio-technical challenges.

However, this detailed account should not be seen as a "one-size-fits-all". My research approach allows pragmatism in meeting the emergent needs for methods and techniques. Therefore, the detailed account of the sessions and what takes place between sessions should be considered as guidance on conducting the study and an adaptive approach that follows a pragmatic strategy in responding to emerging research needs. As a response to such needs, the following are two main differences in the proposed EDPD framework from Hagen et al.'s framework could translate into similar online community research:

Developing a better understanding of the community norms, culture and historical evolution. The literature review and observational study stages added before the start of the PD process have helped identify the historical development and evolution of the community norms, leading to charting the research context and, in turn, an informed design of the online sessions. Figure 3.4 shows an elevated view of my research framework process divided into several phases with a series of sessions grounded in the web-based, distributed study setting, described earlier in 3.1.

As shown in the top and bottom of the process, in parallel with conducting the project, two key activities supplement the process:

- Reaching targeted community groups using relevant communication channels, such as a mailing list or a Facebook group, tools, such as Microsoft Office or a WhatsApp call, and methods, such as online group meetings or social voting.
- Researching and designing suitable collaboration and analysis activities, which this guide unravels throughout its phases.

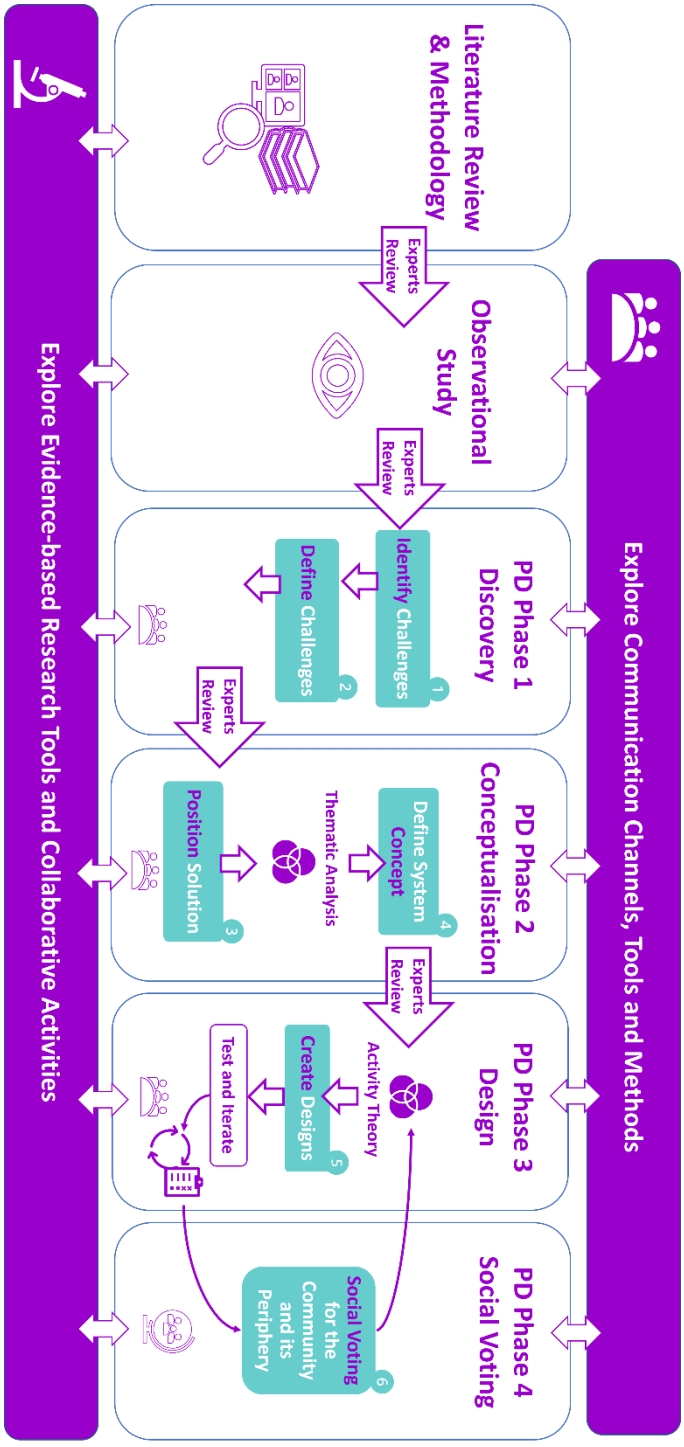


Figure 3.4 Ethnographically-informed Distributed Participatory Design framework process that depicts this research phase.

Adaptive phasic approach. In addition to the literature review and observational study, the EDPD framework is divided into several phases based on the session's themes and analytical approach. It also reshapes the delivery method of sessions using a unique set of online activities, methods and analytical tools depending on the emerging needs and aim of the phase. Further, greater community acceptance includes taking a phasic approach to study the community, as their participation and voices in co-constructing the IT platform are subject to

constant change, owing to the diversity in their visions, local usage cases, and timelines (Huybrechts et al., 2018). This, in turn, intersects with Edwards et al.'s (2007) suggestion that socio-technical infrastructure ensembles cannot be "built" but only "fostered" over time; "they are always constructed in many places, combined and recombined, and they take on new meaning in both different times and spaces" (Edwards et al., 2007, p. 7). Therefore, my research co-design sessions are divided into the following phases:

PD Phase 1 – Discovery (covered in Chapter 5): This phase involves bringing users along with designers to discover the challenges from the community perspective and understand the context of impactful solutions. This phase involves two co-design sessions, "**Identify**" and "**Define**", used as a discovery tool in which methods and techniques can help map the challenges and opportunities and understand the context. In this phase, participants collaborate on grouping and deciding what is relevant for analyses, helping me create outcome summaries that gain everyone's acceptance before moving on to the next session or phase. This helped participants build ownership over the process and confidence in their capabilities for not only producing data but also analysing it.

PD Phase 2 – Conceptualization (covered in Chapter 6): In this phase, participants collaborate on consolidating a system solution concept, bringing focus and defining the research direction and expected outcomes through engaging diverse representation of the community. This phase involves two co-design sessions, "**Position**" and "**Concept**", used as a conceptualisation tool to engage with the community in a process that leads to a short statement outlining the core system concept and the project's scope. This phase may require onboarding new participants to validate the research direction and diversify the voices. It also may witness an accumulation in the data to understand and identify the emerging themes while exploring and complementing the challenges discussed in the Discovery phase. Chapter 6 covers how thematic analysis has helped identify the themes forming the bases of the next phase.

PD Phase 3 – Design & 4 – Social Voting (both covered in Chapter 7): This phase is about transforming the concepts from Phase 2 into prototypes and materialising the design of the system through an iterative process that involves the sessions participants and the global community. This phase involves iterative "**Create**" sessions, including prototyping activities through online virtual board tools or sketching with a pen-and-paper method. It also invites the wider community to comment, vote on other tool design ideas, and add new ideas for all to discuss and vote on using an online social voting tool. If the thematic analysis from Phase 2 uncovers many socio-technical challenges, analytical lenses such as the activity theory can help derive insights and design requirements for the new system.

Therefore, this PhD research does not follow a sequenced process of data collection and analysis leading to findings. Instead, the researchers and participants undergo several data collection and analysis stages until they identify a solution that addresses the challenges. The justification for how each session in this research context was conducted is based on a combination of Hagen et al.'s framework and the outcome of the earlier session or phase. For example, after running three sessions, there was an emergent need in the Conceptualisation phase for onboarding new participants who reviewed the research direction and completed the collected data.

The distributed PD approach used in this PhD research focuses on enriching the framework with online research tools and methods, which is made possible by the advancement of networking and remote collaboration tools. My proposed framework can be used by small teams with the support of a group of researchers, who provide their expertise, review steps, and assistance when needed, especially when moving from one phase to another, as shown in Figure 3.4. Further discussion of the contributions and concepts that underpin these differences is covered in Chapter 8, which looks back on how the research has unfolded. It also provides a short section describing the use of the framework and highlights the key differentiating features, such as the demarcation of the iterative process, to aid potential users in applying and adapting the framework to specific use cases.

3.5 Chapter Summary

The research methodology has undergone many phases of development. However, the core focus is understanding the social-technical challenges and exploring design approaches to support the onboarding process for newcomers, leading to a better retention rate. As discussed herein, after developing a clear picture of the research context using ethnography, the research benefits from the Hagen et al. (2012) framework in building a distributed PD methodology with a set of methods that fit the context and distributed nature of Wikipedia's community.

The method choice is guided by the need to achieve both breadth and depth in the PD approach exercise. For this reason, the study will use synchronous methods that allow deep deliberation of design alternatives and asynchronous methods that require little time commitment and can scale to the wider community.

Different methods serve different purposes in that they:

- Help to identify existing social arrangements, pain points and limitations for design options.

- Support the deliberative process of defining how things ought to be, of creating a vision for the future of Wikipedia training and editing.
- Enable the evaluation of design alternatives.
- Enable quiet, neglected voices to impact the system design.

Chapters 5 to 7 show how this research explores Hagen et al.'s (2012) framework's possibilities and takes a phasic approach to create an environment where Wikipedia trainers and trainees contribute as co-designers following the proposed EDPD framework, which is utilised while focusing on a distributed research approach since my prospective participants' recruitment would be web-based and serve a distributed, global audience.

Therefore, creating an infrastructure for a Wikipedia training platform for collaborative writing involves more than designing computer artefacts. Equally important is the acceptance of the Wikipedia community in sustaining the platform's improvement and utilisation in onboarding new editors. This view has guided this research methodology by pragmatically adapting its method to the emergent needs of Wikipedia's diverse community and opportunities throughout the research timespan, leading to a framework for gaining higher community acceptance of the solution.

Chapter 4 Observational Study

Chapters 2 and 3 help define this research focus and the PD process. The meaningful involvement of the end-users is essential to the success of the PD process (Hagen et al., 2012). However, to facilitate a meaningful discussion, it has been vital for this PhD research to develop detailed accounts of the context in which the proposed design will be placed, the mundane effort involved in conducting training activities, and Wikimedia's innovation practices. The study of the Wikipedia community within their environment and understanding how training is accomplished was essential for designing the sessions that directly involve the end users in the design process. As shown in Figure 4.1, this chapter covers an observational study stage that leads to the start of the sessions covered in later chapters.

This chapter covers an ***Observational Study*** of the current collaboration models in training Wikipedia newcomers, focusing on studying remote Wikipedia training and editing experience to investigate potential barriers. It also aims to explore the innovation process at Wikimedia.

I have already been involved in the community, as a Wikipedian, since 2018; however, for this study, my attendance at four Wikipedia training events was as a passive observer. The data from this observation stage was analysed and used in the sessions covered in Chapters 5 to 7.

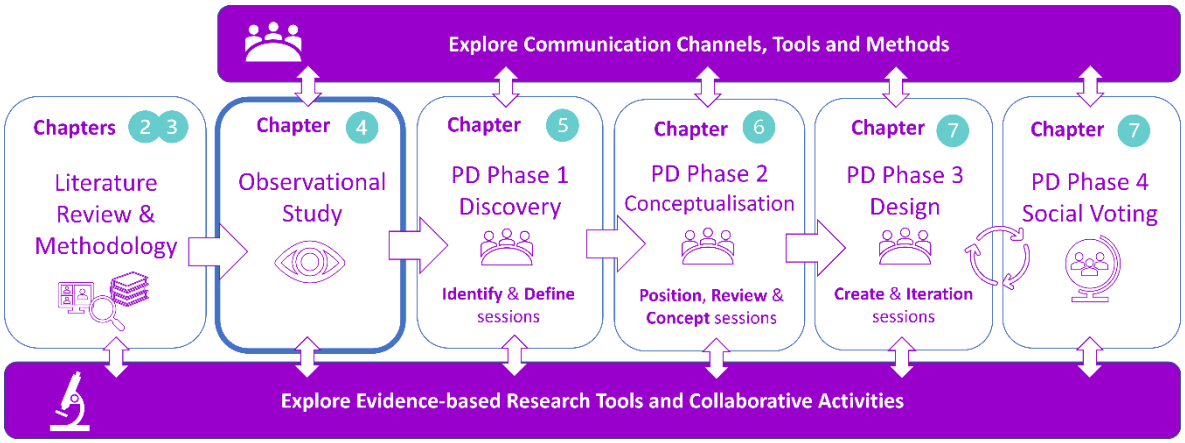


Figure 4.1 The observational study phase is highlighted in relation to other stages of this research.

In-depth observation as part of an ethnography of a community originated from the social science discipline and is known for producing detailed descriptions of the research context (Lazar et al., 2010). However, researchers are encouraged to be careful as it could overwhelm the researcher and cause "difficulties in meshing its work with that of other disciplines"

(Shapiro, 1994, p. 417), which is HCI in this research case. Therefore, since the aim is to conduct an observational study for designing a collaborative tool that a wide range of audiences can use, the conceptual framework shown in Figure 4.2 shows how the data collection and research scope has been limited to exploring specific groups of participants and understanding the context for designing the sessions.

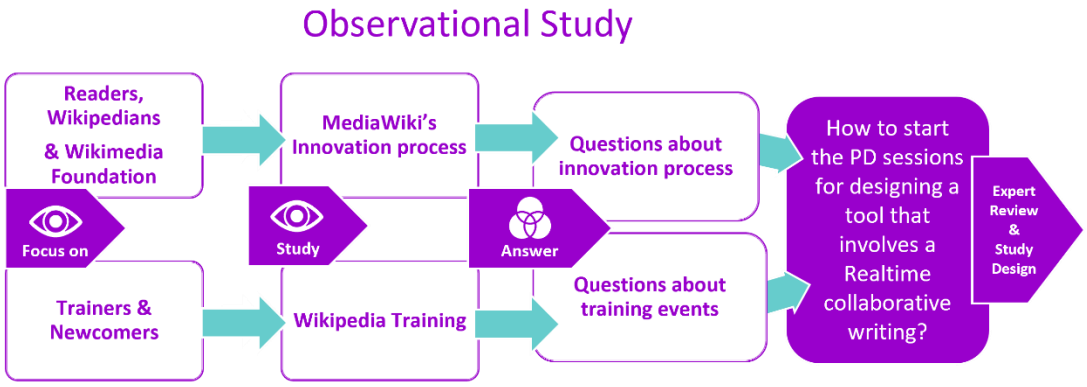


Figure 4.2 Observational Study stages and questions led to the sessions' design.

As shown in Figure 4.2, this chapter answers the following sub-question: **Q1 What are the challenges of introducing real-time collaboration to the Wikipedia community, especially in training events?** and covers contributions C1: and C2:, discussed in Chapter 1**.**

There are many entry points to study Wikipedia's community and its milieu. Figure 4.2 illustrate how I have decided to focus on two community spaces in parallel for the purpose of this study. Therefore, this chapter investigates readers, Wikipedia end-users, Wikipedians and the Wikimedia Foundation's influence on the MediaWiki[42] (the software used on Wikipedia) to answer the following sub-questions about the innovation process:

Q1.1. What processes does Wikimedia have to respond to the user's needs?

Q1.2. What process does Wikimedia use in designing new features?

Q1.3. Why has the real-time collaboration feature not been introduced yet?

Second, the Wikimedia UK's Wikipedia training was recognised as an ever-evolving pragmatic environment to study the possibility of introducing real-time collaboration that lacked research. Therefore, learning more about Wikipedia training through an observational study would help define the following questions about Wikipedia training events:

Q1.4. What is the training activity diagram?

[42] https://www.mediawiki.org/wiki/MediaWiki

Q1.5. What are the collaborative writing modes during training?

Q1.6. What are the challenges and opportunities for newcomers and trainers?

The aim is to shape effective sessions for co-designing a tool that incorporates a real-time collaborative writing mode.

4.1 Introduction

This chapter will focus on describing the research setting using online ethnography to help structure the co-design sessions rather than ideating design decisions or implications.

During my PhD research, there was a need for a pragmatic approach to involving Wikipedia's global community and actively taking part in multiple roles. For this observational study, I needed to distance myself and passively observe the community to look at the bigger picture of the research setting and understand the current collaboration model.

Most research papers that cover design using ethnography describe strategies for conducting co-located research. Virtual ethnography came in response to the need for methods similar to those used in face-to-face interactions in conventional studies of physical workplaces. This shift and the challenges and opportunities associated with it are well documented in Hine's (2000) work. For example, in principles four and five of virtual ethnography, Hine discusses the difficulty in defining the field site or setting boundaries prior to exploration since the community is not located in one physical space. However, to overcome similar challenges, Hine conveys virtual ethnography as an adaptive method that adjusts itself to the circumstance and covers strategies for ethnographers to adapt to the "contemporary Internet" (Hine, 2015). This resembles the strategy followed in adapting the research methods and activities to conduct an online ethnography for Wikipedia training.

4.1.1 Choosing the Research Site

As discussed in Chapter 2, learning to edit Wikipedia individually is a relatively complicated task that may take several paths, albeit primarily leading to the liquid model of collaborative writing or even leaving the community (Li & Farzan, 2018). For example, Li & Farzan discuss how even though Wikipedia training sessions are good in attracting newcomers, the lack of engagement after the Wikipedia training event leads to a poor retention rate. Therefore, the authors invite the Wikipedia community and trainers to benefit from the opportunity to increase the possibility of trainees engaging after training events. This shows that even Wikipedia training requires revisiting its collaboration model.

The community always needs more Wikipedians who understand its norms and effect change that would help Wikipedia meet the demands for adaptation and change to fit into the ever-evolving ecosystem surrounding it (Farzan et al., 2016). This is where training comes in to prepare the newcomers to harmonise their contributions to the community, resulting in less friction owing to a poor understanding of community rules and tools. Therefore, this research focuses on studying Wikipedia training as the research site due to its potential to answer the research question.

4.2 Method

Chapter 3 reviews the growing interest in the methods that study the context in which the system would be used. HCI researchers have developed a greater interest in incorporating ethnographic tools to help develop a deep understanding of the communities and exchange information to help them design for the proper context (Lazar et al., 2010). Ethnography provides unique and useful tools, such as observation for studying the Internet, which develops a deep understanding of how a group of people utilise a set of technologies to constitute the contemporary Internet (Hine, 2015). This does not come without challenges for ethnographers when dealing with a complex environment in which: "the frames of meaning-making that the ethnographer could potentially pursue are initially unpredictable, often diverse, and can require considerable agility of method and mobility of focus to explore". (Hine, 2015, p. 40). Therefore, it is advantageous to look at ethnography as an adaptive method designed to fit the circumstances in which it is implemented.

Section 4.3 covers the method followed in exploring the MediaWiki innovation process, while Section 4.4 discusses the observations from attending four Wikipedia training sessions. Finally, Session 4.5 covers the implications of this study on designing the sessions discussed in the next three chapters.

4.3 Exploring MediaWiki Innovation Process

The routes available to Wikipedia community members or readers to influence or participate in software development are explored in the next section to understand the Wikipedia community norms in participating in the software development process. This would help understand the relations between Wikipedia readers, Wikipedians and the Wikimedia Foundation and its influence on the MediaWiki system design to answer questions about the innovation process shown earlier in the introduction.

Hine argues that using a "combination of strategies helps to develop insights into the multiple meanings of a single phenomenon." (Hine, 2015, p. 89). Therefore, this section covers

my strategy to collect data by going through the real-life experience of Wikipedia community members who would like to participate in Wikimedia's innovation process. This presents how the MediaWiki system takes input from and engages users in the design process from the perspective of Wikipedia readers, peripheral community members, and experienced Wikipedia editors. According to Given (2008, p. 489), this exploration approach, which is referred to as "lived experience", gives "a representation and understanding of a researcher or research subject's human experiences, choices, and options and how those factors influence one's perception of knowledge."

While Chapter 2 explores introducing the real-time collaboration that came to a halt in 2018, this section covers Wikimedia's innovation process that responds to user needs. Also explored is the new users' and expert community members' involvement in the technical evolution direction of Wikipedia and its sister projects. This exploration has produced several paths that influence changes in MediaWiki's socio-technical systems design, proposing features such as adding a real-time collaboration mode, which is discussed below. However, other routes, such as design decisions based on analysing the usage metrics of Wikipedia, are not covered due to time and access limitations.

The following sections aim to draw a picture that conveys the challenges someone could go through and help address the questions about the innovation process outlined at the start of the chapter.

In addition to answering the questions, this exploration demonstrates how the change in this platform is challenging, not only because of part of the core community's resistance to change but also because of the lack of organisation and complicated user experience in the innovation process. This extends the contribution of this thesis by covering why real-time collaboration is not introduced and how to introduce it through a new systematic development process that could be used in the future.

My previous experience as a community member has helped identify several avenues for user innovation in Wikipedia. However, in this section, I attempt to investigate the end-user experience with the innovation process regardless of my know-how. I, therefore, have set aside my prior knowledge to ask how a reader or editor would propose a new feature and participate in the innovation process.

4.3.1 Requesting a New Feature Through a Google Search

A regular reader of Wikipedia or a casual contributor (peripheral community member) wants to suggest functionality that enhances her reading or editing experience on Wikipedia. To

achieve that, googling “Wikipedia feature request” would result in a Wikipedia category page with a collection of ideas made by Wikipedia community members (Wikipedia Feature Request - Google Search, 2023), shown in Figure 4.3. The category[43] page states, “this category may be a bit pointless, since the developers will probably never see it. To post a feature request, please visit Wikipedia:Bug reports and feature requests or Wikipedia:Village pump (proposals).” (*Wikipedia Feature Requests - Wikipedia*, 2022).

The page was created in 2007, with few updates, having the last one in 2019. It is an attempt by the community to collect feature requests scattered on several Wikipedia pages, such as the link to this page: “Wikipedia: Feature request (archive)[44]“. As with many other pages, all direct users to the “bug reports and feature requests” Wiki page shown in Figure 4.4.

Figure 4.3 Wikipedia: Feature request webpage.

It is worth noting that both pages, “Wikipedia:Bug reports and feature requests” and “Wikipedia requests for new software features”, as shown in Figure 4.3, would redirect to the same article, which is illustrated in Figure 4.4 below.

[43] https://en.wikipedia.org/wiki/Help:Category

Figure 4.4 Wikipedia's bug reports and feature requests.

As shown, the article is filled with technical terms that direct users to Wikimedia's Phabricator,[45] which is Wikimedia's software suite of tools intended for reporting bugs and project management. It also mixes the bug report and feature request as shown in the following, which proceeds the guidelines of reporting a bug or requesting a feature: "You've got a bug or feature request, and you're itching to report it." The guidelines take the user to a page titled "How to Report a Bug" and then a section explaining the process for reporting a bug or new feature. This leads the user to Figure 4.5, which shows the Phabricator's home page, where the third option is creating a new feature.

This path requires the user to create a MediaWiki account. After creating one, a guideline step asks the user who would like to request a feature to search for an existing request on the Phabricator before proceeding. This is sensible but requires dealing with technical terms and the Phabricator formatting and markup, mainly known to experienced Wikipedians.

[45] https://phabricator.wikimedia.org/

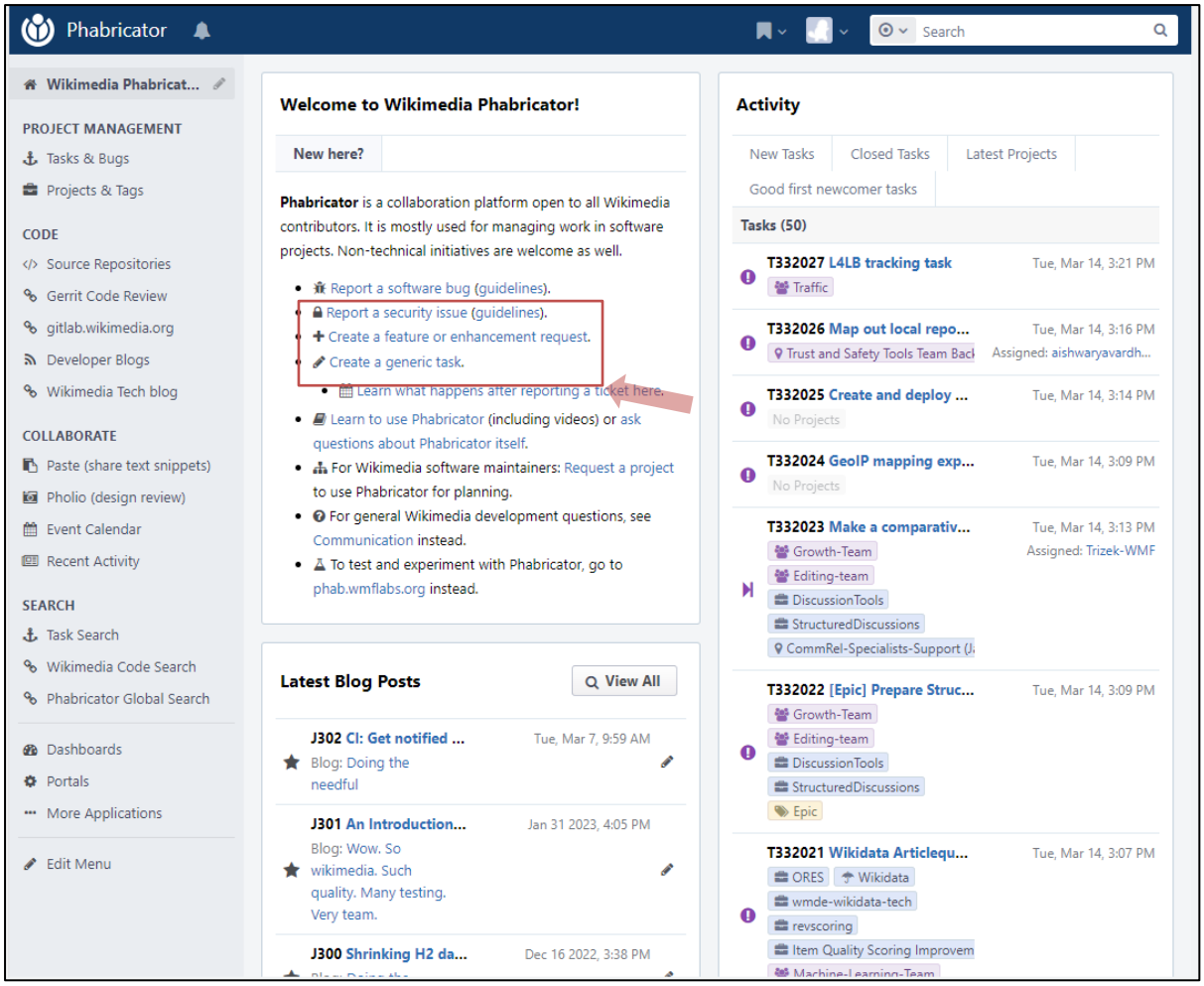


Figure 4.5 Phabricator home page.

The third item in the welcome section of the top left leads a user to the page shown in Figure 4.6 to submit her feature request.

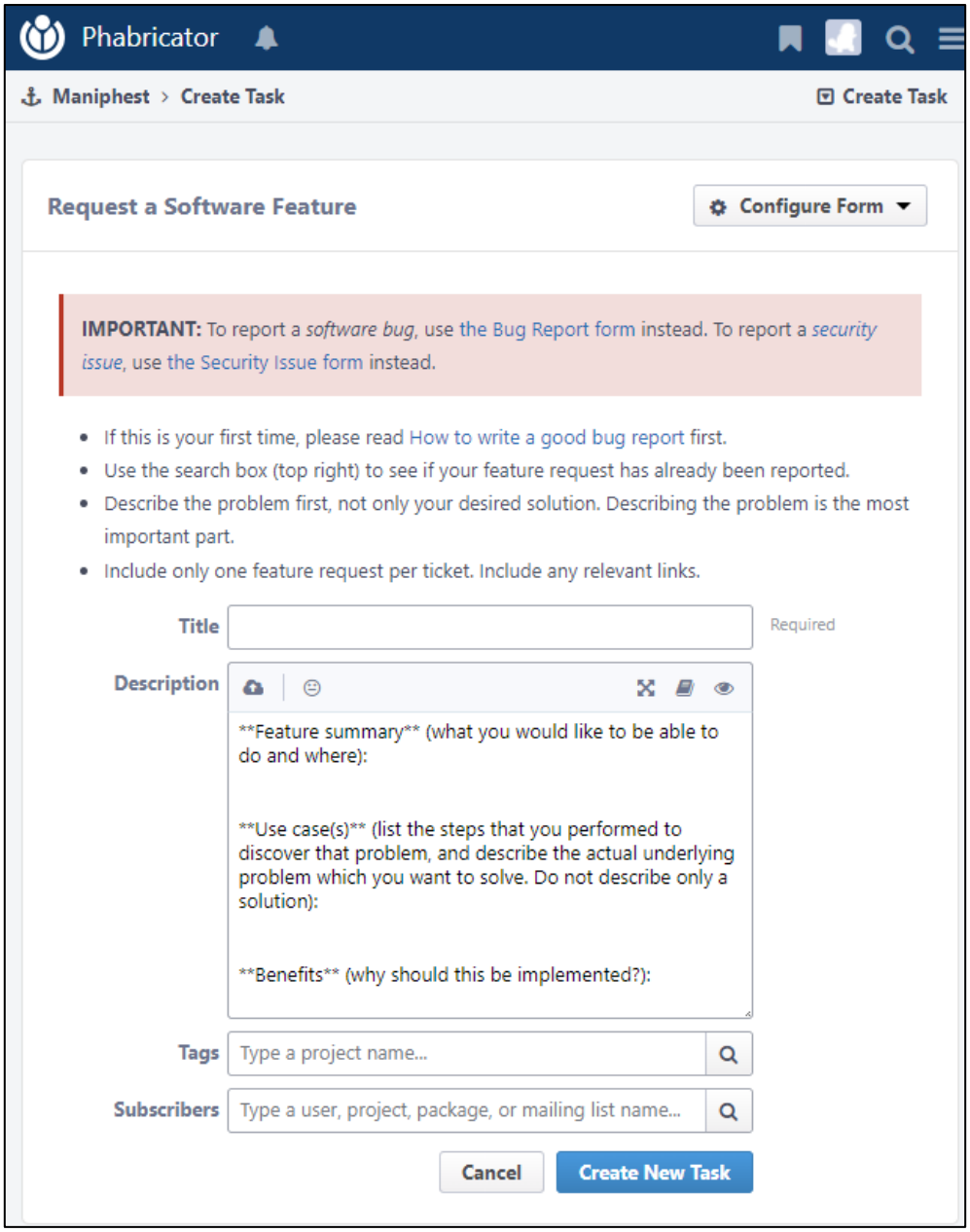


Figure 4.6 Requesting a feature on the Phabricator interface.

4.3.2 Proposals in the Village Pump

Another method to request a feature is the Village Pump[46], where the Wikipedia community discusses their policies, ideas, and technical needs, communicates with the Wikimedia Foundation, and more. As a term, the Village Pump could vary across Wikipedia versions and cultures, which shows how tricky it is to standardise practices such as requesting features across language versions. For example, it is "Diwan" in the Urdu language, which translates as "court"

[46] https://en.wikipedia.org/wiki/Wikipedia:Village_pump

in English, while it is “The Bistro” in the French version, and “Maidan” in Arabic, meaning “an open space” or “town square”.

Engaging in this space requires a medium to advanced level of expertise in using the MediaWiki software editor. For example, to reply to someone, one needs to pay attention to the indention level and add a colon at the start of the sentence to indent her contribution under the original message. It also requires enough time to enter discussions that could take a long time before reaching a consensus or getting a reply from the Wikimedia Foundation development team. *Wikipedia:Village Pump (WMF) - Wikipedia*, (2022) is an example of a lengthy discussion between the community and Wikimedia over problems with mobile editing.

4.3.3 Submitting Requests to the Community Wishlist Survey

Another path for users’ influence on Wikimedia projects is orchestrated by the Wikimedia Foundation’s Community Tech team, which operates the Community Wishlist Survey (CWS).[47] Annually since 2015, the team has been releasing a cross-project survey that appears for almost 14 days on Wikipedia pages, as shown in Figure 4.7, inviting contributors to suggest features and enhancements they would like to see in all language versions of Wikipedia and its sister projects, such as Wikibooks and Commons. A proposal could be declined for several reasons. For example, it may be refused when the complexity level or technical need for the features is beyond the capacity of the Community Tech team, when the feature could not be delivered in one year, or when it had been declined before or was already in Community Tech teams’ plans.

Figure 4.7 Cross-project survey ad asking the community to suggest and vote on features (Community Wishlist Survey, 2022).

[47] https://meta.wikimedia.org/wiki/Community_Wishlist_Survey_2022

The Community Teach team was formed by the Wikimedia Foundation upon a request from the core community of contributors to respond collaboratively to the community of users' needs for creating or modifying the existing bots, gadgets, Wikimedia extensions, and admin tools used by the "core community" of Wikimedia projects' active contributors (Mediawiki et al., 2022). It collaborates with Wikimedia's volunteer and staff developers from other teams on small projects that directly impact the community of editors in all languages and could be delivered quickly. For example, one of their projects is adding an autosave functionality to the Wikipedia article editor.

For a feature proposal to be successful, it should undergo three phases. The process starts by asking users to submit proposals in any language to the Wishlist survey, though the team states that English is preferred to facilitate discussions between the team and other editors. The user is expected to submit the proposal in a specific format that covers the following points:

- The problem they want the feature or enhancement to focus on.
- The targeted audience: whether it is the community of editors, administrators, Wikimedia commons project users, or Wikipedia users.
- How the problem is being dealt with now.
- What is their proposed solution?
- Is there a similar endeavour being undertaken in the Phabricator?

In the second phase, the Community Tech team compiles the proposals, asks for further information on missing details, and merges similar requests. The third phase involves voting on the proposals. As regards the Community Wishlist Survey 2015, out of 103 ideas, only the top 10 submissions with the highest votes reached the team's backlog of projects for further research and work. Figure 4.8 shows the categories of the requested features between 2015 and 2019, showing both accepted and rejected proposals.

Category	Proposals in top 10					Most votes for a proposal
	2015	2016	2017	2019	Total	
Admins and stewards	–	1	0	–	1	86
Admins and patrollers	–	–	–	1	1	157
Anti-harassment	–	–	0	1	1	118
Bots and gadgets	2	2	2	1	7	111
Categories	1	–	–	0	1	84
Citations	–	0	0	1	1	108
Editing	1	1	3	2	7	127
Maps	–	–	–	0	0	59
Miscellaneous	1	4	2	0	7	154
Mobile and apps	–	0	0	0	0	55
Moderation tools Moderation and admin tools (2015)	0	1	–	–	1	84
Multimedia and Commons combined in 2017	1	0	1	0	2	94
Notifications	0	–	–	1	1	101
Programs and events	–	0	1	0	1	111
Reading	0	0	0	1	1	130
Search	0	0	1	0	1	110
Special pages	0	–	–	–	0	22
Talk pages	1	0	0	–	1	66
Templates	1	–	–	–	1	87
Watchlists	2	0	0	1	3	104
Wikidata	0	0	0	0	0	68
WikiProjects	–	1	–	–	1	72
Wikisource	0	0	0	1	1	113
Wikiversity	0	–	–	–	0	4
Wiktionary	–	0	0	0	0	40

Figure 4.8 The categories of the requested features in the Community Wishlist between 2015 and 2019. (Community Wishlist Survey Categories - Meta, 2021).

Real-time Collaboration in Wikimedia's Wishlist

The pursuit of a real-time group collaboration feature that allows everyone editing Wikipedia to collaborate similarly through tools, such as Etherpad, appeared in the 2016 Wishlist (*Real Time Group Editing Proposal - Community Wishlist Survey*, 2016) shown in Figure 4.9. It was ranked 183 out of 265 proposals.

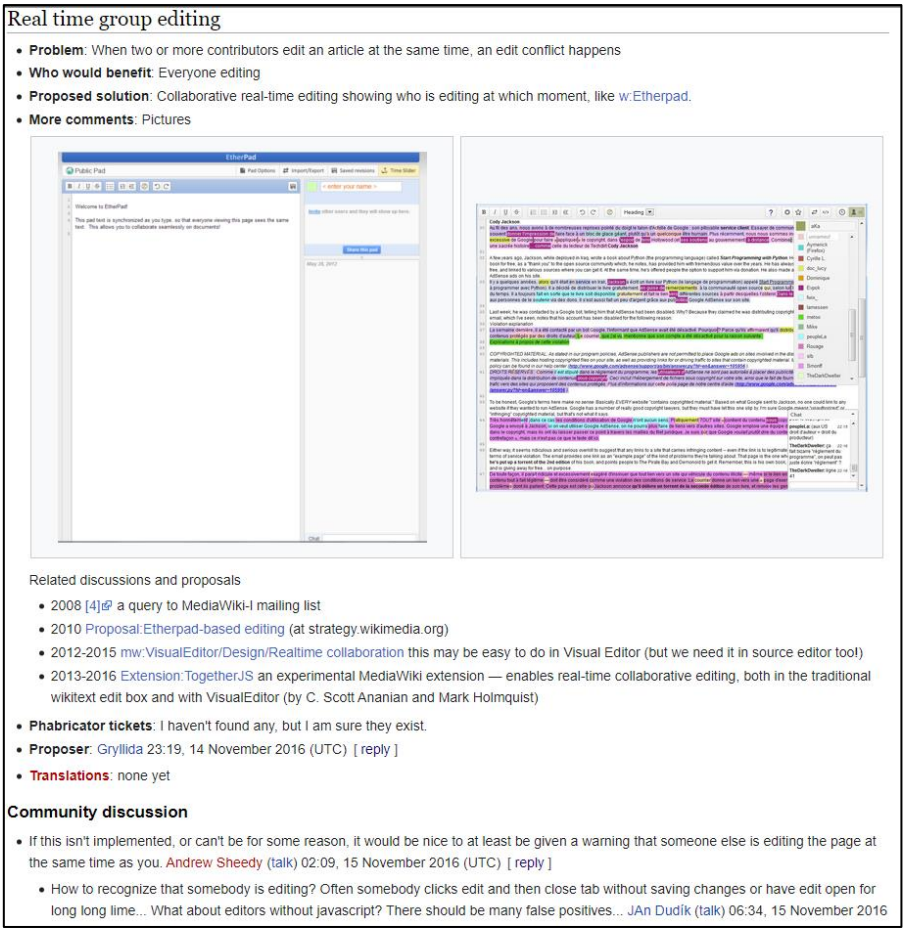


Figure 4.9 Real time Group Editing Proposal - Community Wishlist.

Furthermore, it was opposed by ten community members, supported by seven, and one neutral. Many opposed the proposal for several reasons, some of which are:

Some community members noted that this feature would only support those with JavaScript enabled on their web browser. This would exclude other editors who have JavaScript disabled and prefer using the plain Wikitext[48] editor over the visual one, creating more edit conflicts with other editors with JS disabled.

1) One user describes this as too advanced technology for Wikipedia users.
2) Some Wikipedia community members predicted that the feature would bring fundamental change that is too complex to be implemented by the community tech team in one year.

[48] "Visual Editor is a "visual" way of editing Wikipedia. Editing Wikipedia has historically required people to learn wiki markup [Wikitext using a Wiki Source editor] a fairly complex markup language, even to make tiny changes to a page." (*VisualEditor/Portal - MediaWiki*, 2022) https://en.wikipedia.org/wiki/Help:Wikitext

3) One user voted to oppose the proposed idea of changing such a core functionality of Wikipedia for all users but supported the many benefits that using real-time could bring to team efforts at WikiProject.

Those who supported the proposal endorsed using real-time collaboration in specific contexts rather than across Wikipedia. For example, they proposed using it in workshops to draft more neutral articles with fewer edit conflicts, bringing more quality and efficiency. It was also mentioned that such a collaborative editing feature is a vital prerequisite in similar contemporary online work environments of the "modern world"; and that it is unfortunate not to see it implemented in Wikis.

This request for real-time collaboration was mentioned again in the 2020 Wishlist Survey (*Real Time Group Editing Proposal 2 - Community Wishlist Survey*, 2020). It demanded real-time collaborative features, such as editing in the Wikiversity Project,[49] to help the community provide a more interactive method for learners, similar to the ones on other websites, such as Udemy[50] and Moodle.[51] However, the proposal did not make the final list, as the community tech team indicated it needed to be more specific and marked it as "Outside the scope[52] of Community Tech".

4.3.4 Wikipedia Users Volunteering at the Product Teams

Wikimedia has a team of designers across the organisation's projects. The teams' page welcomes contributors to get involved through volunteering to "Review, create or contribute to Phabricator design tasks".[53] Also, Wikipedia has formed a "Core Experiences" group, which encompasses the Web, Growth and Editing Product teams, where each team has created a space for discussion about their work in the talk pages of their projects or by directing them to the Phabricator. For example, the Editing Team is responsible for editing tools across Wikimedia. It welcomes feedback about MediaWiki Visual Editor through a dedicated feedback page with the latest talk page features and a discussion page that redirects to the same feedback page.[54] The Growth team[55] introduces features to increase newcomers' engagement and growth in the retention and activation of new editors. That team is currently experimenting with a new homepage for new users, including information on how to edit and direct access to mentoring

[49] https://en.wikiversity.org/wiki/Wikiversity:Main_Page
[50] https://www.udemy.com/
[51] https://moodle.org/
[52] https://meta.wikimedia.org/wiki/Community_Tech#Scope
[53] https://www.mediawiki.org/wiki/Design
[54] https://www.mediawiki.org/wiki/Talk:VisualEditor/Feedback
[55] https://www.mediawiki.org/wiki/Growth

support. There are several ways to contribute to the growth team outlined in their engagement guide,[56] ranging from sharing an experience to contributing to the Phabricator.

4.3.5 Working at Wikimedia Foundation as a Staff Member

The governance of Wikipedia's decision-making process is considered one of the researchers' interest areas to better understand self-governed online communities. Rijshouwer et al. (2021) mention that Wikipedia has undergone several stages of development. The most recent of which led Wikimedia to define formal roles, such as the Community and Developer Liaisons, to channel the community needs, reaching the Wikimedia Foundation's objectives in serving the community. These roles usually demand deeper knowledge of Wikipedia community practices and long history of volunteer contributions. Going through the background of Wikimedia staff, one would find many examples of volunteer community members who have transitioned from admins or experienced contributors to paid staff members. The effect of this on product design is hard to measure.

4.3.6 Contributing to the Wikimedia Strategy

Between July 2009 and July 2010, community discussions and proposals shaped the movement's five-year strategic plans. For example, proposals for real-time editing (*Collaborative Google Wave-Based Wikipedia Editing - Strategic Planning*, 2009; *Etherpad-Based Editing Proposal - Strategic Planning*, 2010) have made their way to the discussion, leading to Wikimedia's product map, as shown in Figure 4.10. One featured proposal did not mention real-time collaboration, though. Still, it was referenced as an abstract proposal by those focused on the real-time feature and was referred to as the "possible way of implementing the more abstract proposal" titled "Improve software, keep up with the times" (*Improve Software, Keep up with the Times Proposal - Strategic Planning*, 2009).

[56] https://www.mediawiki.org/wiki/Growth/Engage_with_the_Growth_team

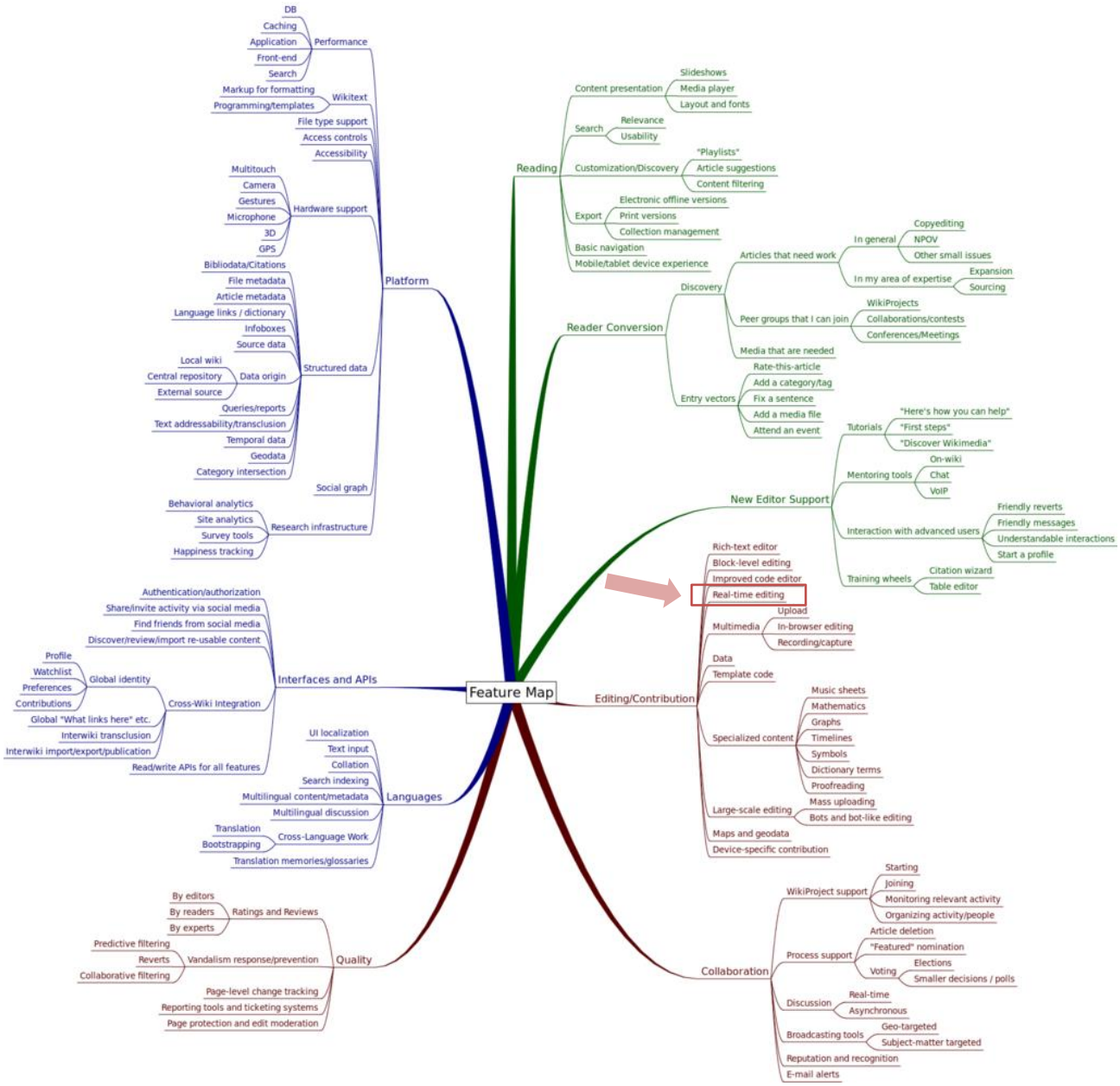


Figure 4.10 Wikimedia's product map.

As discussed in Section 3.1, one of Wikimedia's methods for user innovation is direct consultation with the community (*Talk Pages Consultation 2019 - MediaWiki*, 2019). Recently, Wikimedia has involved the community in a participatory process to develop the 2030 Movement Strategy (*Movement Strategy/About - Meta*, 2020). The primary participants are the organised groups connected with the Wikimedia Foundation and diverse groups of contributors.[57] The method also included what Wikimedia calls "New voices", such as readers, experts and partners (*Wikimedia Strategy Participants*, 2017). The involvement, as illustrated in Figure 4.11, is composed of a number of phases. The process resulted in a recommendation

[57] https://meta.wikimedia.org/wiki/Movement_Strategy/About

that influenced areas such as improving the user experience (*Wikimedia Strategy on Improving the User Experience*, 2018) and ensuring equity in decision-making (*Wikimedia Strategy on Improving Equity*, 2018).

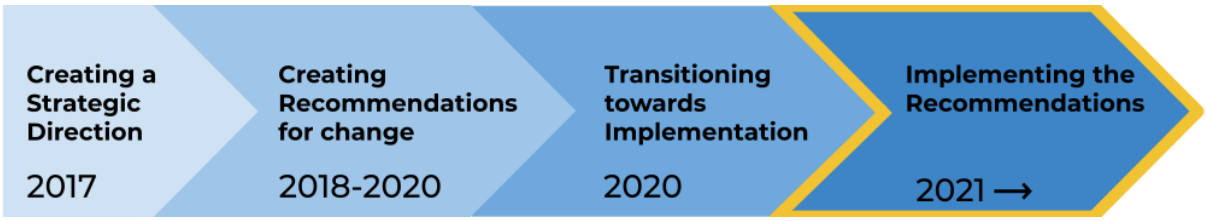


Figure 4.11 Wikimedia strategy process (Movement Strategy/About - Meta, 2020).

4.3.7 Discussing MediaWiki's Innovation Process

The following sections discuss key observations to find answers to the innovation process questions posed in the introduction section. Insights highlighted in the earlier section are reflected on, which would later lay the foundation for answering the innovation process questions.

The Affordance of Involving the Community in the Innovation Process. Incorporating new major features in Open-Source Software societies usually involves the end user in requirement elicitation, software development, and decision-making. Most communities favour advanced technological alternatives, even if the cost or risk of overpromising is too high (Androutsellis-Theotokis et al., 2010). After the proposition and voting, however, in most cases, the feature selection for implementation becomes the responsibility of the core team of developers.

Most Wikimedia project development processes follow the same pattern, slowly catching up with the technological and software engineering trends to advance the platforms with reasonable cost and the engagement of their community.

Wikimedia Foundation is trying to preserve the characteristics of an Open-Source Software development process while pushing on its strategic plans through its large team of full-time employees, developers and designers. Their team is responsible for maintaining over 20 projects, with some, such as Wikipedia, being available in over 300 languages. Therefore, finding the balance between what the community needs and what the Wikimedia Foundation could do is tricky.

Innovation Process. Reviewing most of the paths for innovation in Wikimedia, it is recognised that the focus is on the core community perspective on product development. As a result, the experience of newcomers is dictated by those with a particular and advanced use of the skills that enable them to contribute to Wikimedia's projects and tools, such as the Phabricator.

Therefore, Wikimedia's innovation process should be more diverse, involving a more PD approach at all phases of product development, from identifying the problem to the development and iterating for use. In addition, new alternative flow models should be in place to solve long-lasting issues or bring a feature that might cause significant socio-technical outside-inside change to the platform. However, as will be discussed in Chapter 7, the size of the community involved in each stage should be approached cautiously, as the involvement of the wider community at an immature level might overwhelm staff and the community.

Contemporary Online Work Environments. As discussed earlier, some community members call for adapting to contemporary technological advances. This is not limited to real-time collaboration; it requires updating the processes developed to engage their broader audience in shaping the products. Benchmarking Wikimedia to recently launched similar projects, one could find interesting insights. For example, WikiJS launched in 2016 a new open-source project built in JavaScript to "Make documentation a joy to write using Wiki.js's beautiful and intuitive interface." [58] WikiJS has over 18k community stars on GitHub compared to over 3k for MediaWiki. This, however, is an unfair comparison since the MediaWiki project on GitHub is just a mirror of Wikimedia's open-source version of Github. Nevertheless, WikiJS has the benefit of using Github's discussion feature to engage one of the biggest developers' communities in discussing bugs and voting on their importance, as shown in Figure 4.12.[59]

[58] https://js.wiki/
[59] https://github.com/requarks/wiki/discussions

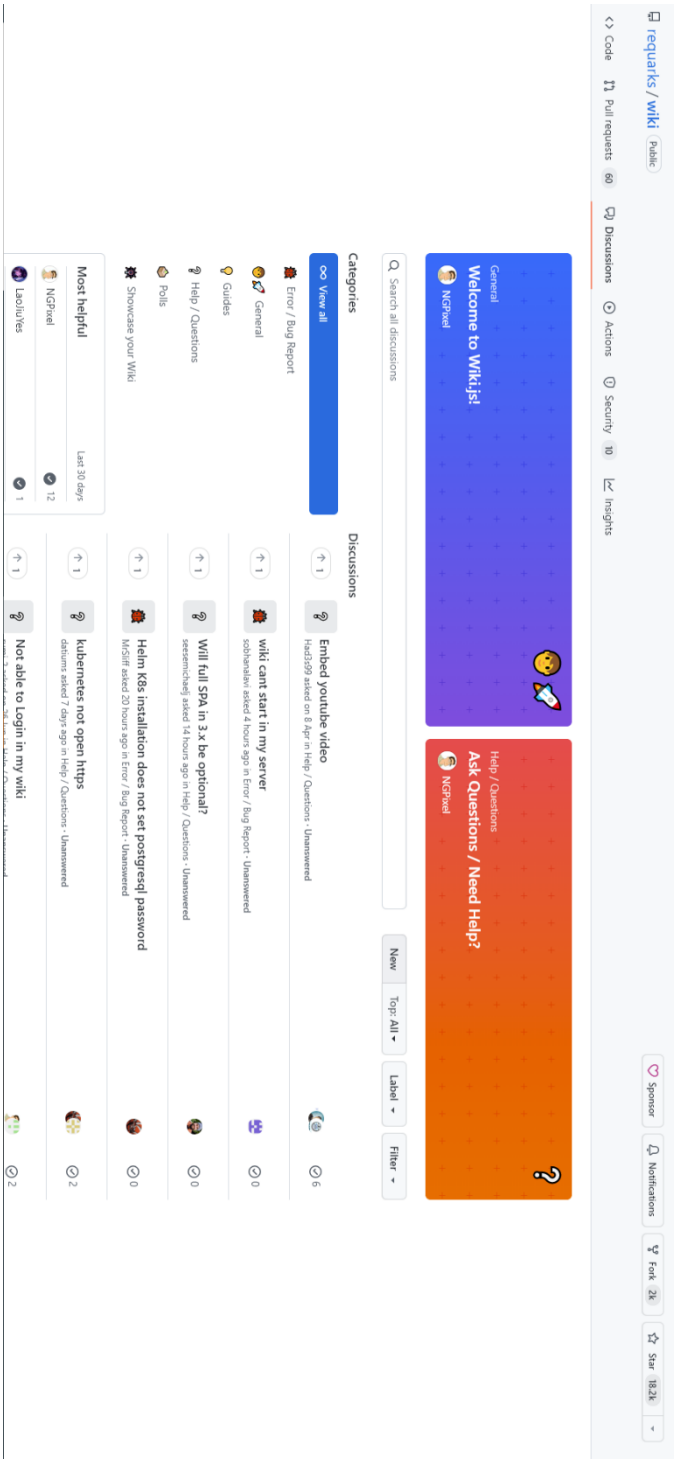


Figure 4.12 WikiJS repository on GitHub.

In addition, WikiJS uses a commercial interactive service called “Canny” for new feature requests, shown in Figure 4.13.[60] This service helps communities have a centralised place for discussions around new features, prioritise work, and eliminate replicated requests. Moreover, the end user gets notified of their requests’ progress.

[60] https://requarks.canny.io/wiki

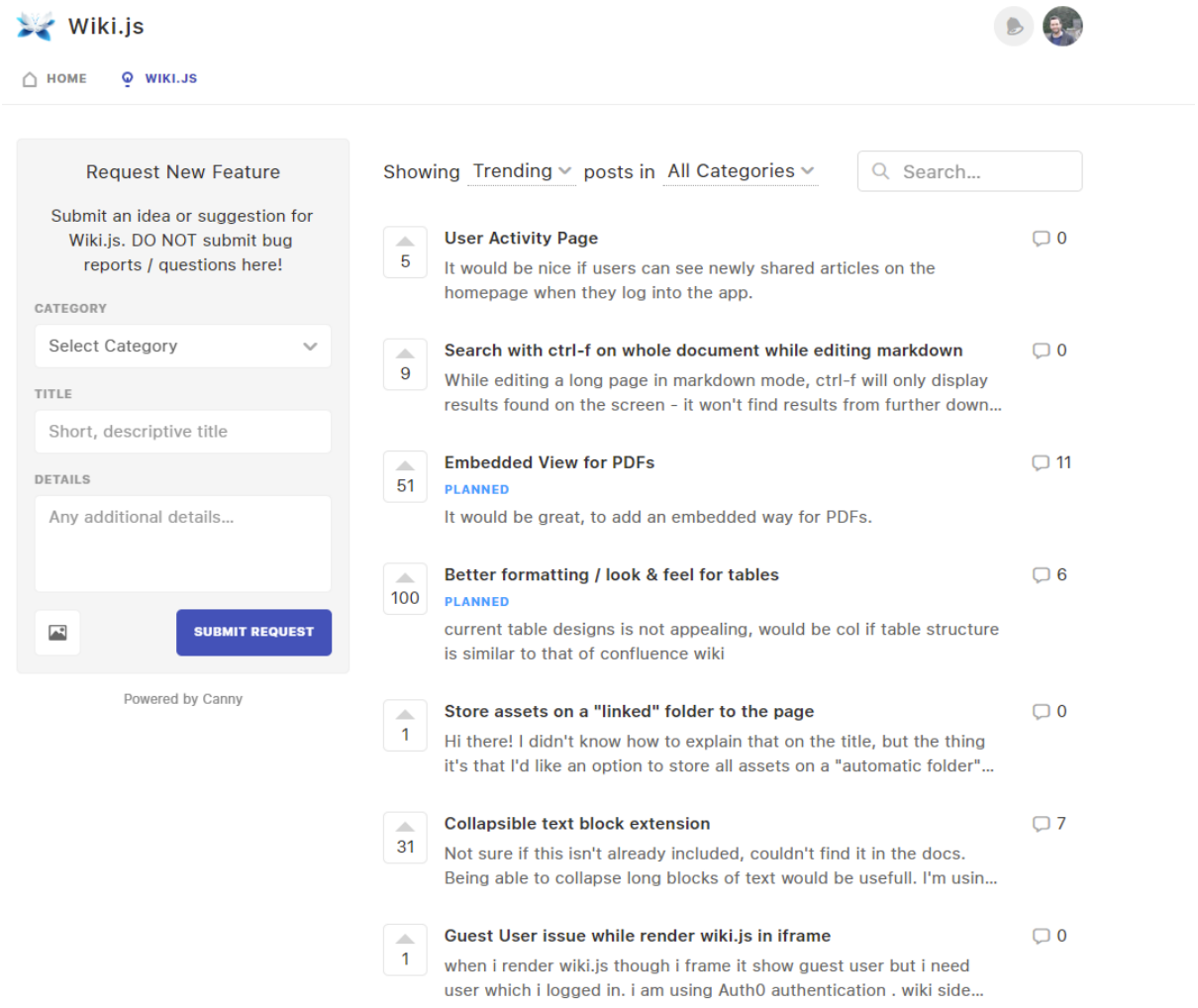


Figure 4.13 WikiJS repository on Canny.io.

Looking at this perspective shows the importance of involving the broader community in this research project's sessions (discussed in Chapters 5 to 7).

Technical Know-how. Most routes to end-user engagement in software development demand the technical knowledge of an above-average Wikipedia user who would be able to navigate through, understand, and deal with terms or pages, such as "Bug", "Village Pump", or the Phabricator system.

One of the paths many users could take is finding the page in Figure 4.3, which has a link to "Wikipedia:Bug Reports and Feature Requests". The "reporting a bug" terminology is not a common term to many people outside of the tech domain, and mixing the path to the process of filing a bug and requesting a feature might diminish the importance of getting the community requests for features. Furthermore, inside the bug reports and feature requests help page, which explains the process of asking to fix a bug or a feature, many terms are only understood by core

contributors to Wikipedia, which would exclude a large category of readers and peripheral community members.

Global Involvement of the New Voices. Wikimedia Foundation is a global movement engaging a broad spectrum of voices in its development. This is apparent in their strategy manifesto, which has involved the broader audience in the movement direction, including readers, experts, and partners. However, the involvement of the new voices has decreased with moving up in the process of developing and implementing recommendations, with more focus turned to Wikipedia's community of contributors. Additionally, the strategy text is translated into 30 languages having limited pages with various levels of coverage and paths to discussing and contributing or suggesting features. This could be attributed to limited resources compared to the scope of its projects, which is beyond the reach of this research.

The Politics of Community- Wikimedia Relation. Politics in the core community and Wikimedia relationship creates tensions in the innovation process. For example, the more Wikimedia improves the editing experience and innovation process, the more newcomers may challenge the core community. To this point, Ford & Wajcman (2017, p. 521) discuss the following:

> "Although the Wikimedia Foundation invests significant resources into improving (and simplifying) the user experience, they have faced significant pushback from the older community of (mostly male) users, who are unhappy with the influx of those they regard as inexperienced editors entering the system. Indeed, the conservatism of this early Wikipedia community has been identified as the primary cause of Wikipedia's decline in editor numbers."

Therefore, Wikimedia's approach to introducing change and new features faces contradicting views and requires finding balance. The WMF sometimes face criticism for putting resources into paying consultancy and contractors to help with its strategy or to bring marginalised people, who, according to some, produce lower quality contributions, instead of investing the money in supporting the existing community. This can be found in recent discussions of the Wikimedia Movement 2017 strategy.

4.3.8 Answering Innovation Process Questions

This section builds on previous discussions to answer the innovation process questions.

Q1.1 What processes does Wikimedia have to respond to the user's needs?

Wikimedia project users have no central online space for people to propose features and provide feedback on the MediaWiki design or features. Even though Phabricator acts as the central place where most product development paths lead, many users do not have the technical know-how to request or vote on features there. In addition, each Wikipedia version has its methods or dedicated pages for contributing to software development, increasing complications for Wikimedia in responding to rapidly changing user needs.

Q1.2 What process does Wikimedia use in designing new features?

The Wikimedia Foundation is a non-profit established in 2003 to provide "the essential infrastructure for free knowledge." They "host Wikipedia, the free online encyclopaedia, created, edited, and verified by volunteers around the world, as well as many other vital community projects." (*Wikimedia Foundation Website*, 2020). Since 2003, the design process has developed significantly to adapt to the new market and user needs. However, Wikimedia's approach to designing features and engaging the end users has not kept pace with this field's development for many reasons, some of which were discussed earlier. In addition, this is an under-researched area as most research focuses on Wikipedia's lack of technological development rather than product development and the innovation process.

Wikimedia maintained the robustness and fixed the inherited issues of the Wiki software. They focused on the technical side of its projects, the Engineering Design that, according to Löwgren (1995), relies more on well-documented problems to fix. However, Wikimedia's focus shifted recently to include many processes and end-user engagement options, supporting the socio-technical needs of their community and facilitating more Creative Design that relies on "understanding the problem as much as the resulting artifact." (Löwgren, 1995, p. 521).

In 2016, a collaborative process was launched to define a statement of purpose for its design team due to the lack of a "clearly defined or shared understanding of the purpose for design at the WMF." (*Team Practices Group*, 2022). They brought stakeholders internal to the non-profit and external from the Wikimedia community to:

> "Articulate the shared aspirational vision of design at the Wikimedia Foundation, to articulate what it is at a high level the group is attempting to achieve. In addition to providing a clear vision, this document will aid in making decisions for the design group, such as those in regards to (but not

limited to) staffing, organising, implementing projects, and strategising." (*Design Statement of Purpose - MediaWiki*, 2016)

Since its establishment, its research team has conducted plenty of generative and evaluative studies to meet the foundation's mission of serving the community and its audience of readers. Furthermore, after receiving funding in 2009, the foundation established dedicated research for the development of the platform, including qualitative studies to understand how barriers can be removed for readers to participate in Wikipedia (Vora et al., 2010).

Q1.3 Why has the real-time collaboration feature not been introduced yet?

Evidently, the effort to introduce real-time collaboration was stopped mainly due to the social complexities outlined in greater detail in Chapter 2. Hence, this question has been partially answered in the second chapter, where Wikipedia's attempts to include real-time collaborative editing are reviewed. Wikimedia Foundation's research team does not lack the expertise that could potentially experiment with methodologies to answer this research question of how to introduce real-time collaboration. This begs the question of why it has not been introduced yet. The answer comes from the exploration covered earlier in this chapter, which shows how users' needs are addressed and the different routes for requesting and prioritising features in MediaWiki.

Kittur et al. (2006) recommend constant monitoring of the shifts in user base needs by collaborative knowledge systems designers. Once the collaboration community is stable, attracting a high influx of new users, it should shift its focus from powering experienced users with new features to providing novice users with a better user experience (A. Kittur et al. 2006). Now, the Wikimedia Foundation has many staff members and volunteer technical teams working on projects that tackle various problems and needs in their open-source software. Their medium-term plan for 2019 acknowledges the need to improve the tools and documentation to reduce the barriers and complexities of its technology stack, making the experience more enjoyable for the new editors and technical contributors to Wikipedia and its sister projects. In addition, they aimed for a 5% increase in retention rate for technical community members who make remarkable efforts to understand active community needs (*Wikimedia Foundation Medium-Term Plan 2019/Platform Evolution - Meta*, 2019).

However, most methods used by end-users in steering Wikipedia's infrastructure, such as proposing a feature on the Village Pump (proposal) Wiki page,[61] submitting a bug report, or

[61] https://en.wikipedia.org/wiki/Wikipedia:Village_pump_(proposals)

filing a feature request on the Phabricator,[62] favour addressing the core community's suggestions. The new editors and average Wikipedia users have minimal options to propose, advocate, or vote on the major changes, such as adding real-time collaboration to this global platform. Most of them would not have the time, Wiki know-how or ability to find the early discussed paths; even if they do, it is unlikely to comprehend the community terminology, processes and norms. This puts the most significant stake in advancing the platform in the hands of the core community, which has resisted changing the status quo of asynchronous editing at Wikipedia, as seen in the proposal earlier. The resistance, however, should be seen as "transformative, rather than destructive because it is a common feature of socio-technical change." (Hyysalo et al., 2016, p. 29). Hyysalo et al. draw on the work of Kline (2002), who describes several cases where the technology proposed to the community as a way to modernise was initially refused, but later adopted after they were altered to accommodate their needs, and sometimes, serving a purpose other than the initial intended use. This indicates the importance of studying communities based on the outcome of their innovation process and understanding these communities' needs and motivation to participate in the process (Heiskanen et al., 2010). This will be the focus of the observational study in Section 4.4, which will study the Wikipedia training and should help design an online environment for different, even resisting, voices to come together to answer this research questions in the sessions covered in Chapters 5 to 7.

4.4 Wikipedia Training Observation

This section covers the process and findings from observing 4 Wikipedia training sessions involving 56 trainers, organisers and trainees, with some sessions having the same trainers. Later in Section 4.4.4, the key findings are reflected on. But first, the following Section 4.4.1 will attempt to clarify some of the ambiguity that could result from the different terms used in referring to Wikipedia training events.

4.4.1 What Is a Wikipedia Training Event?

The terminology for Wikipedia training events is evolving. For instance, two trainers may use different titles from the same type of event. However, when organisations such as Wikimedia UK mentions "Wikipedia training events", they usually refer to teaching newcomers how to edit Wikipedia and give them an experimental asynchronous editing session to experiment with what they learn with a trainer's close support and supervision. Another known title for Wikipedia training events is Edit-a-Thon.[63] This title is usually associated with

[62] https://en.wikipedia.org/wiki/Wikipedia:Bug_reports_and_feature_requests
[63] https://en.wikipedia.org/wiki/Edit-a-thon

events that commemorate a certain annual celebration or support an activism campaign and its editing phase. Edit-a-Thon gives attendees a quick run through the how-to edit "bitesize training" and much more time for editing, which could last for multiple days. In all these events, preparing an exciting writing worklist (the articles to be written) and sharing it with the Wikipedia community would attract experienced editors or another trainer, who occasionally attend these events to help if the trainer expects many newcomers. Recently, some of these events have been advertised as "Wiki Workshops", which could be less daunting than Edit-a-Thons.

There are subtle but important distinctions between Wikipedia training sessions and "edit-a-thons". For example, most training session advertisements aim to motivate newcomers to develop new skills through editing, which leads to social impact, while the edit-a-thon focuses more on bringing more contributions to a specific social cause which is contingent on learning how to edit Wikipedia for newcomers. Such difference in structure leads to a difference in motives of well-established Wikipedians joining these events. Experienced Wikipedians mostly join the training sessions to help the trainees, but they mostly join edit-a-thons to promote, as a community, certain editing agendas for a specific area on Wikipedia, which may or may not involve helping newcomers. Also, edit-a-thons may be run by a motivated organiser from a Wikipedia user group or non-affiliated community member about a topic without a trainer or affiliation with a Wikimedia chapter. In contrast, the training session necessitates having a trainer who, for the most part, would have a connection with a local chapter of the Wikimedia Foundation.

In my thesis, I focus on collaborating with Wikimedia UK to explore "Wikipedia Training" as the broader case for research.

4.4.2 The Observation Method

The observation method involves data collection from the fieldwork, ranging from the large-scale survey and conducting hundreds of interviews "to the lone researcher recording information collected through participant observation in a small-scale case-study." (Marshall, 1994, p. 182). As discussed earlier, the focus has been on Wikipedia training. After attending a three-day training by Wikimedia UK charity, I became an accredited trainer who is highly engaged in the community. Following running several training sessions, I attended the Wikimedia UK in Education summit,[64] helping me to meet numerous prominent community

[64] https://wikiedusummit.coventry.domains/

members. This has led to my involvement in helping with the transition to online training, considering the relevance of my research topic to stakeholders.

With assistance from collaborators at Wikimedia UK, I was given access to training sessions to collect observation notes to study remote Wikipedia training and online collaborative Wikipedia editing processes. This has helped me take a closer look at the needs of the trainers and participants with diverse demographic backgrounds to answer the questions about the training events posed by this research question about real-time collaboration.

Having experienced the complexity of running an event, I realised that adding a research element to an already demanding event for the trainers, organisers, and attendees is not an easy ask. Therefore, I have minimised my interference in the event, explaining the benefit my research could have on improving future training.

I digitised the participants' information sheet and consent forms, which were ready to be signed electronically. During the observed online sessions, I introduced myself to the group, turned my camera off, and took notes until the concluding remarks, when I thanked all for having me and left my contact details for further questions or inquiries.

Notes were taken anonymously about participants' individual and collaborative editing and training strategies; this did not include their Wikipedia edits. Fieldnotes were collected using word editing software and later moved to the Miro[65] board (an interactive virtual whiteboard collaboration platform for distributed teams) for analysis.

T1. First Training[66]

Thirteen participants joined training between 12 and 4 pm on the 29th of January, 2021, using a licensed Zoom account to edit articles about women in art. It included staff members, an organiser, a speaker and experienced Wikipedians interested in the topic. Before the training, the advertisement had mentioned that laptops or PCs are preferable to tablets or mobile phones. It also encouraged anyone to register, as no previous editing experience was needed. It assured the participants that many breaks would take place and Zoom breakout rooms would be available to avoid "Zoom-fatigue".

Participants were asked to:

- Create a new Wikipedia account if one does not have one already.
- Register at Wikimedia's Outreach Dashboard [67] using a username to track one's edits.

[65] https://miro.com/
[66] The letter "T" in the numbering refers to the training number, and the letter "O" refers to the number of the observation.
[67] Tool which helps with the management of wiki programs and events https://outreachdashboard.wmflabs.org/

- Conduct research in advance.

This was the trainer's second online training session; all previous sessions took place face-to-face. The trainer requested support from the mailing list of other trainers, and one co-trainer was interested in providing technical support throughout the event.

First Training Main Observations:

T1.O1. For participants to commit to the 3-4 hours of training, they need time flexibility:

T1.O1.1. The organiser is covering for another person who could not attend owing to a "pandemic-related childcare situation".

T1.O1.2. A participant said she/he must leave earlier to collect a child from school. "So, if I disappear, don't feel offended."

T1.O1.3. As participants doubted their ability to write articles in the training time frame, the trainer told a trainee in an encouraging tone not to worry about the size of the new articles, stating they should not be "that long".

T1.O2. The training process involved pretraining activities and clear instructions on how to edit or create articles. A relaxing and flexible environment has led to collaborations among participants on findings, resources and worklists:

T1.O2.1. The trainer accessed the event 15 minutes earlier and started receiving participants after 5 minutes of coordinating with the organiser.

T1.O2.2. The trainer paused occasionally for people to ask questions, creating a relaxing environment by encouraging people to feel comfortable and savour some snacks.

T1.O2.3. The trainees discussed the distribution of articles to be created. One participant brought a long list of women to write articles about; another asked if they, too, could write about some of them.

T1.O2.4. The organiser described the worklist, which was similar to the one shown in Figure 4.14, and asked participants to collaborate on the articles and the resources.

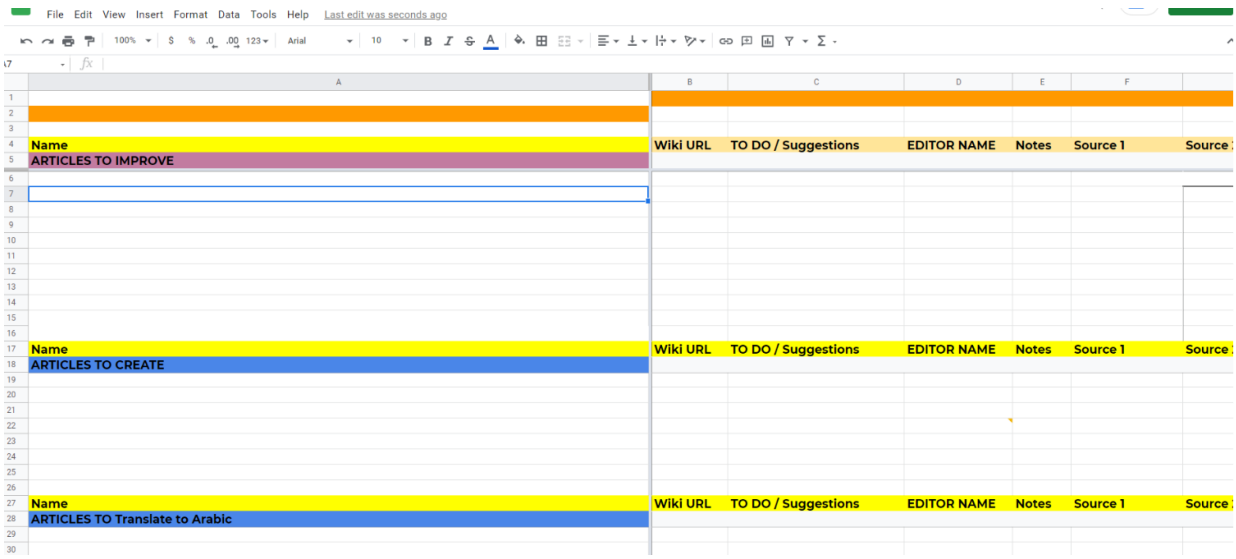


Figure 4.14 Wikipedia training session Worklist.

T1.O2.5. The trainer explained to the participants that they must look at an existing page and then draft something accordingly. Once the draft is done, a participant would get help in putting the formatting together.

T1.O2.6. Apart from writing articles, participants were asked to improve articles through simple tasks, such as adding Wiki links, info boxes, and images.

T1.O2.7. A feeling of appreciation existed for those who developed the working list, and the person who created it was happy to see the articles move from being created to being enhanced in real-time. However, some participants were surprised by the people who were missing from Wikipedia and should be added to the names list, which led some of them to assure the rest that they would do some follow-up editing.

T1.O3. The Training Challenges:

T1.O3.1. After editing, having heard a trainee say, “It’s a bit of a mishmash what I have done”, the trainer intervened and asked the trainee to share her/his screen, although they did not know how to do it.

T1.O3.2. One participant said in a worried but chuckling voice that she was taking the time of the training by sharing the screen and getting help with publishing. Another participant trying to break the moment’s awkwardness, said, “this is so fun!”.

T1.O3.3. Participants lost track of what the trainer did when moving an article from the sandbox to the namespace. However, seeing the trainer do it again, they were impressed by her/his dexterity (It felt as if they were saying it would take me a long time to do such an expert task).

T1.O3.4. The trainer forgot to confirm[68] the participants' accounts,[69] so participants could publish an article after creating an account.

T1.O3.5. The trainer had to explain how to generate the Reference list twice.

T1.O3.6. One trainee was worried after interrupting someone to ask, saying, "Sorry, I am not very good at knowing when I should speak on these things." (Referring to the Zoom application).

T1.O3.7. The trainer mentioned to participants that editing Wikipedia is easy; it is like editing Microsoft Office. However, the rules behind Wikipedia are what people find challenging.

T1.O3.8. Other Technical Challenges

- Some participants left their microphones unmuted by mistake.
- A trainer asked a co-trainer to see if the audio worked and requested her/him to turn up the volume. The co-trainer responded that she/he "have sound issues; going to update Zoom and come back in".

T1.O4. The Training Tool's Functionalities:

T1.O4.1. Splitting people into groups using breakout rooms.

The trainer described how, depending on the number of attendees, the participants would be split into two groups, one for experienced editors and one for those less experienced or who need help. The trainer also mentioned that they might separate, as this could help contain noise in the training session, given the large number of trainees. According to the trainer, the other benefit is that providing help in a separate room would be easier.

Breakout rooms eventually did not materialise, as no one asked for them, and the trainer did not follow up.

T1.O4.2. The organiser informed the participants they would take a screenshot of the Zoom meeting for documenting purposes.

T1.O4.3. Participants were already introducing themselves in the chat as more were joining and could not see the previous chat.

T1.O4.4. People were using chat to ask, and the trainer answered with audio.

[68] https://en.wikipedia.org/wiki/Wikipedia:User_access_levels

[69] On English Wikipedia, only those autoconfirmed/confirmed users can create articles or move articles from the draft space to the main space. Therefore, some trainers tend to elevate the access level for their trainees' Wikipedia accounts to allow them to create new articles. Otherwise, the trainee needs to have a 4-day old account and have made at least 10 edits for it to be auto-confirmed. This rule is different from other Wikipedia language versions.

T1.O4.5. The trainer and organiser mentioned that there is no pressure to keep the camera on. One participant replied, “It’s easier to keep my camera off because of my ADHD. Sorry!”

T1.O4.6. The trainer used Google Slides without Full Screen mode, showing previous and upcoming slides on the sidebar for other participants.

T1.O4.7. The trainer said she/he could share the slides with participants. However, she/he warned trainees that they might be overwhelmed as they had more details not covered at the training, but keep them for reference.

T1.O5. Collaboration and Asynchronous Conflict:

The trainer stated the following to describe the possibility of edit conflict with the asynchronous model of editing:

> “One thing we need to check is column D in the Google Sheet [shown in Figure 4.14 earlier] with articles to edit, where we put the editor’s name. Wiki does not do well because you cannot have two people editing the same page at the same time. It creates edit conflict. So, once you pick the article, put your name, so we don’t end up with duplicated work or edit conflict.”

T2. Second Training

Sixteen participants joined training from 1 to 3 pm on the 11th of March, 2021, using a licensed Zoom account to edit articles about women archaeologists. It included an organiser and experienced Wikipedians interested in the topic.

Second Training Main Observations:

T2.O1. Participants heavily relied on real-time chat for socialising and collaboration:

Similar to the first session, the chatting system in this session was used for the round of introductions, sharing resources, and asking and answering questions.

T2.O2. Collaborating on worklists and sandboxes:

T2.O2.1. The organiser asked if the participants had prepared resources about women archaeologists for creating articles. One participant said, “I was not very active before the session and did not bring a specific name for creating an article”. The organiser told the participants that someone had extra names for editing that could be shared.

T2.O2.2. The trainer was asked about creating a new page, which led them to resume screen sharing to explain how to create a draft in the sandbox. Later, a participant inquired about copying and pasting content for the sandbox to a Wikipedia Article.

T2.O3. Wikitext and the Visual Editor confusion:

T2.O3.1. The trainer explained how to switch to the visual editor, if the source code editing loaded on the first use. Later, one participant said that "I think I have come off the visual editor as my sandbox is just Wikitext; how do I get back to it?"

T2.O3.2. The trainer had to switch to Wikitext editor to explain editing categories.

T2.O4. Wikipedia Functionalities:

The trainer used the alternative sequence from my first observation to explain Wikipedia's different functionalities.

T2.O5. The training outcome:

The lack of records about historical figures frustrated some, although the participants were delighted with the training and expressed excitement about doing more editing.

T3. Third Training

Sixteen participants joined training from 10 am to 1 pm on the 25th of March, 2021, using a licensed Zoom account to edit articles about librarians. It included an organiser and interested experienced Wikipedians.

Third Training Main Observations:

T3.O1. Beyond real-time chat:

One member had a discussion around a question they asked using only real-time chat, without turning on the microphone, even though the trainer replied through voice in real-time.

T3.O2. The training process and varying needs:

T3.O2.1. The training started with explaining the usage of Zoom using Google Slides.

T3.O2.2. Someone asked about creating multiple sandboxes for drafting different articles, which was not mentioned in the training session.

T3.O2.3. The trainer used a publicly accessible Etherpad document (see Etherpad in Figure 4.15 used to coordinate work among the participants).

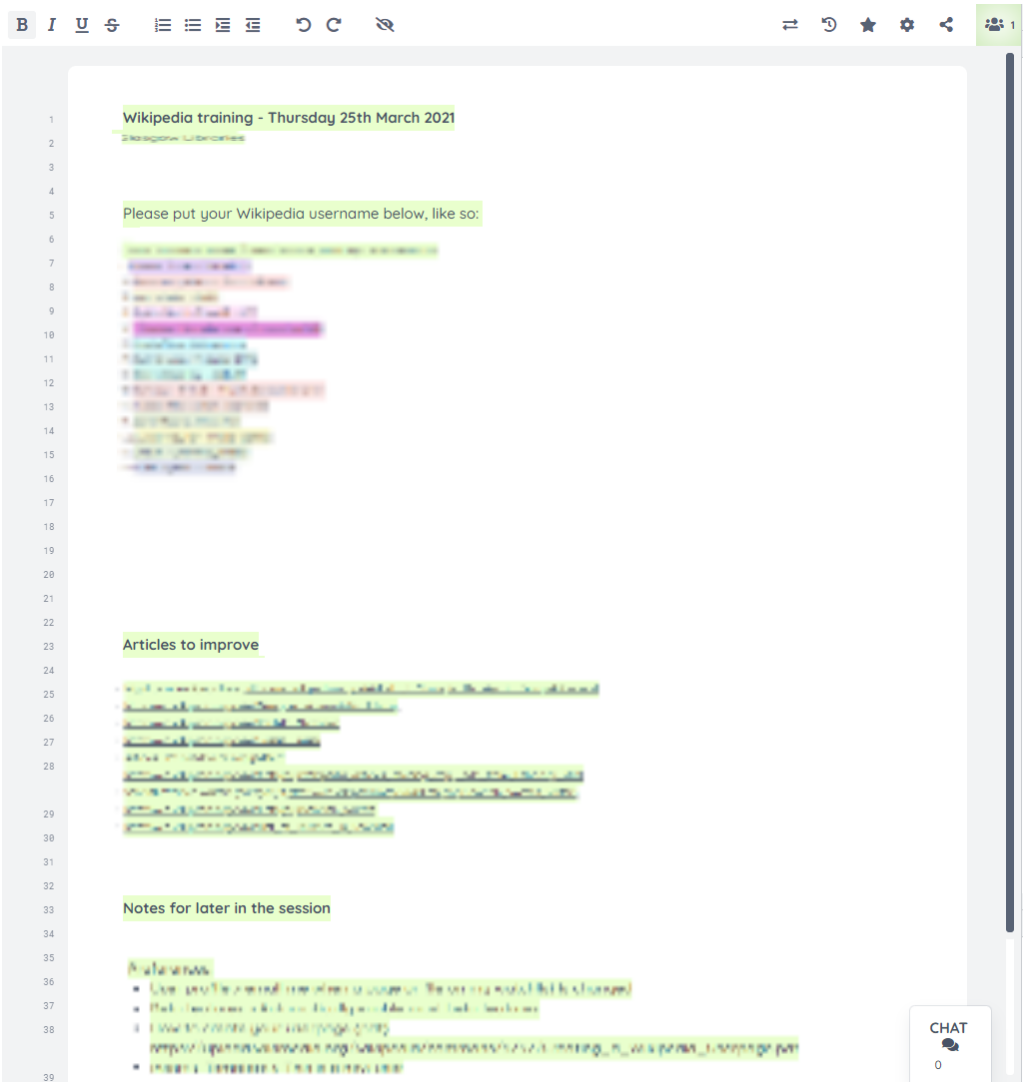


Figure 4.15 Etherpad doc used to coordinate Training 1.

T3.O2.4. The trainer used slides that had screenshots of Wikipedia with red triangles to point out important sections.

T3.O2.5. Some users accessed the training from their mobile phones and PC simultaneously—the phone being for the call and the PC for Wikipedia-related work.

T3.O2.6. The trainer had a training account to show the participants a clean interface similar to their interface. However, the trainer could not demonstrate the publishing stage, as her/his account was intended just for training.

T3.O3. Online training vs face-to-face:

T3.O3.1. The trainer encouraged participants to use their cameras to make the training similar to the face-to-face training session.

T3.O3.2. The trainer used the alternative sequence of slides and examples from the trainer in my first two observations to explain Wikipedia's different functionalities, even though both are colleagues in Wikimedia UK.

T4. Fourth Training

Eleven participants joined training from 1 to 4 pm on the 1st of April, 2021, using Microsoft Teams to edit articles about notable women. It included an organiser and experienced Wikipedians interested in the effort.

Fourth Training Main Observations:

T4.O1. Confusion with source and visual editing.

T4.O1.1. As in almost all training sessions, the participants were confused by switching between the editing modes. For example, one felt it was standard procedure to see the markup code and asked the following, "I am editing, am I supposed to see the markup code within the text? I am having a difficult time ignoring this and being able to parse the text I am editing."

T4.O1.2. Participants were more comfortable using the raising hands feature in the conference call to ask questions compared to earlier training.

T4.O1.3. The trainer relied on the screen-sharing feature to demonstrate to and lecture the trainees.

T4.O1.4. One of the participants asked if the trainer would share the slides after the training.

4.4.3 Analysis

Ospina-Pinillos et al. (2019), who used Hagen et al.'s (2012) framework, intertwined their data collection process with analysis using codes based on Valdez et al.'s (2012) Culturally-Informed Design Framework, shown in Figure 4.16. This has helped them shape stages according to findings on each one, rather than waiting for the end of data collection from all sessions.

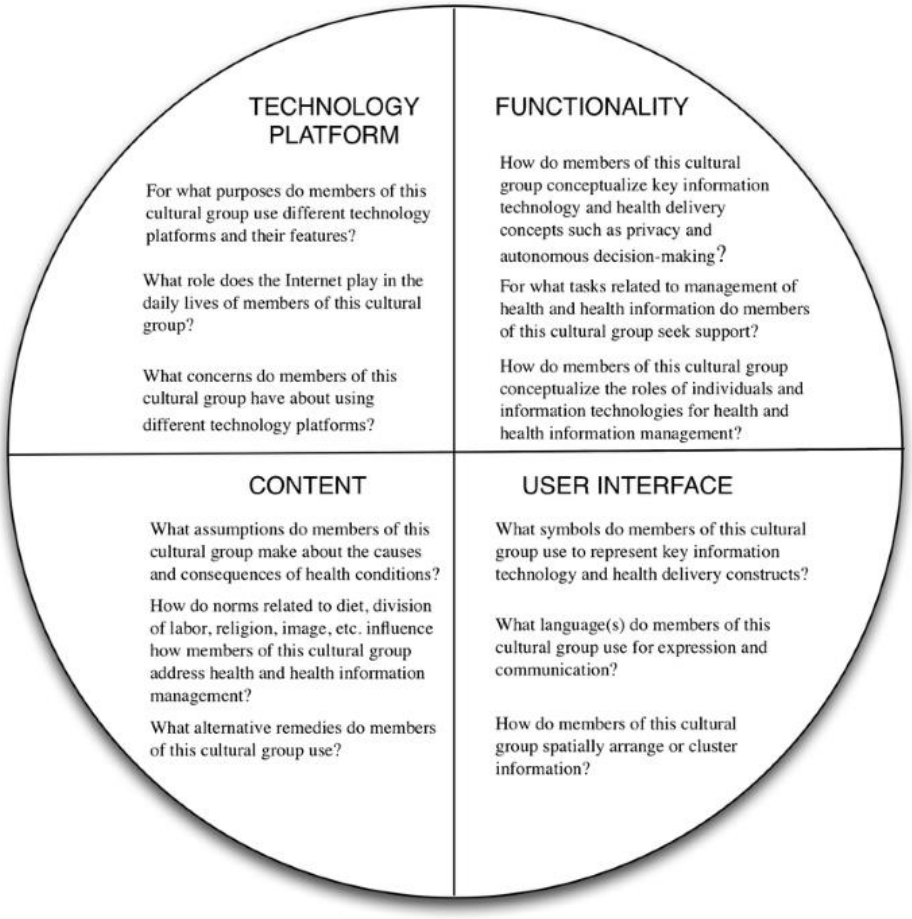


Figure 4.16 Culturally-Informed Design Framework by Valdez et al. (2012).

The authors' deductive coding was complemented with codes from their previous research (Ospina-Pinillos et al., 2018). For this PhD research, after experimenting with the deductive approach using the same framework codes: Technology platform, content, user interface, and functionality, some new codes, such as time, fun, issues, low confidence, and process, emerged that did not fit the deductive approach to the coding scheme, and I had to take an inductive approach to code the data.

The observation notes for each event were moved to sticky notes on a Miro board, later tagged with the date, the role of the person connected to the note, and the code. Figure 4.17 shows a part of the board that covers the analysis of the first session. An adapted intersected circles version from Valdez et al.'s (2012) framework helped find connections between codes. For example, the trainer on the Zoom platform had the challenge of knowing if the participants came back for the break since most of them turned their cameras off. This observation with the challenge code falls under the User Interface, Functionality, and Technology Platform areas. Another example of coding an observation is when the trainer was observed to show users they

could rename themselves using Zoom to facilitate more collaboration. This observation had a collaboration code and fell under the four category-areas.

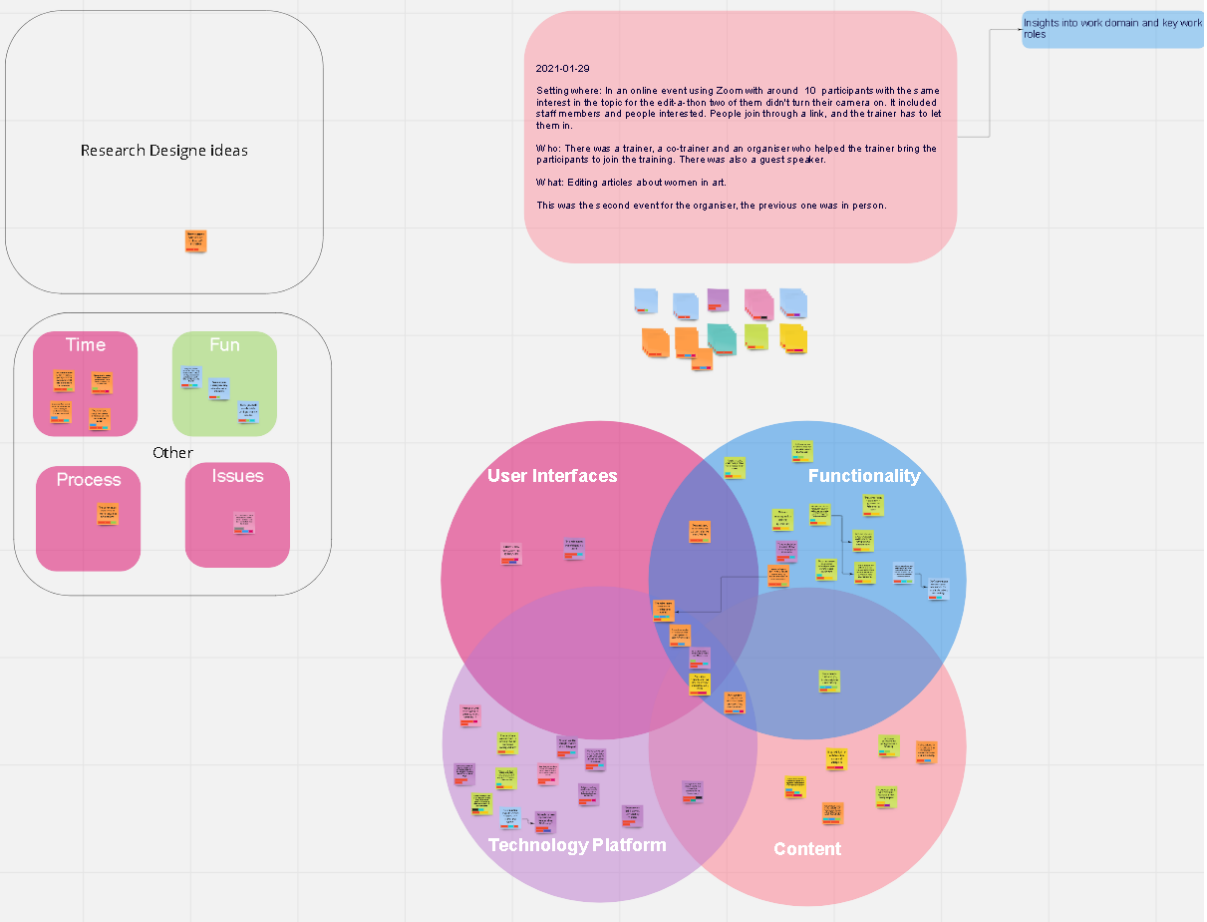


Figure 4.17 The first session section of the Miro board used the analysis of the observational notes.

4.4.4 Answering Questions about Wikipedia Training

Returning to the questions about the study design, this section builds on previous discussions in answering the questions.

Q1.4 What is the training activity diagram?

Based on the data collected from the training events, the first outcome was Figure 4.18, which shows the flow model of virtual training and the different systems used by trainers, co-trainers, organisers, and participants. After arranging the online training with the trainer, the event would be advertised using Eventbrite or other channels. The event takes place on a conference call using a platform such as Zoom, depending on the organiser's preferred tool, the trainer shares a screen with a presentation service, usually Google Slides. Before and during the event, different actors interact with several systems. The trainer sets up the dashboard, and the participants are asked to join using their newly created accounts to track their edits, sometime

before the training but mainly during it, when the trainers dropped the real-time chat PDFs about Creating a Wikipedia User Page,[70] changing preferences, and adding an info box. The trainers sometimes use external tools to demonstrate a point. For example, in one of the training sessions, a trainer shared Denelezh's Gender Gap Tool[71] in Wikidata to illustrate the importance of editing Wikipedia and contributing to filling the gaps. Later, there would be a worklist on Google Sheets or Etherpad, where participants coordinate their tasks.

[70] https://upload.wikimedia.org/wikipedia/commons/5/57/Creating_a_Wikipedia_Userpage.pdf
[71] https://web.archive.org/web/20220706132050/https://denelezh.wmcloud.org/gender-gap/

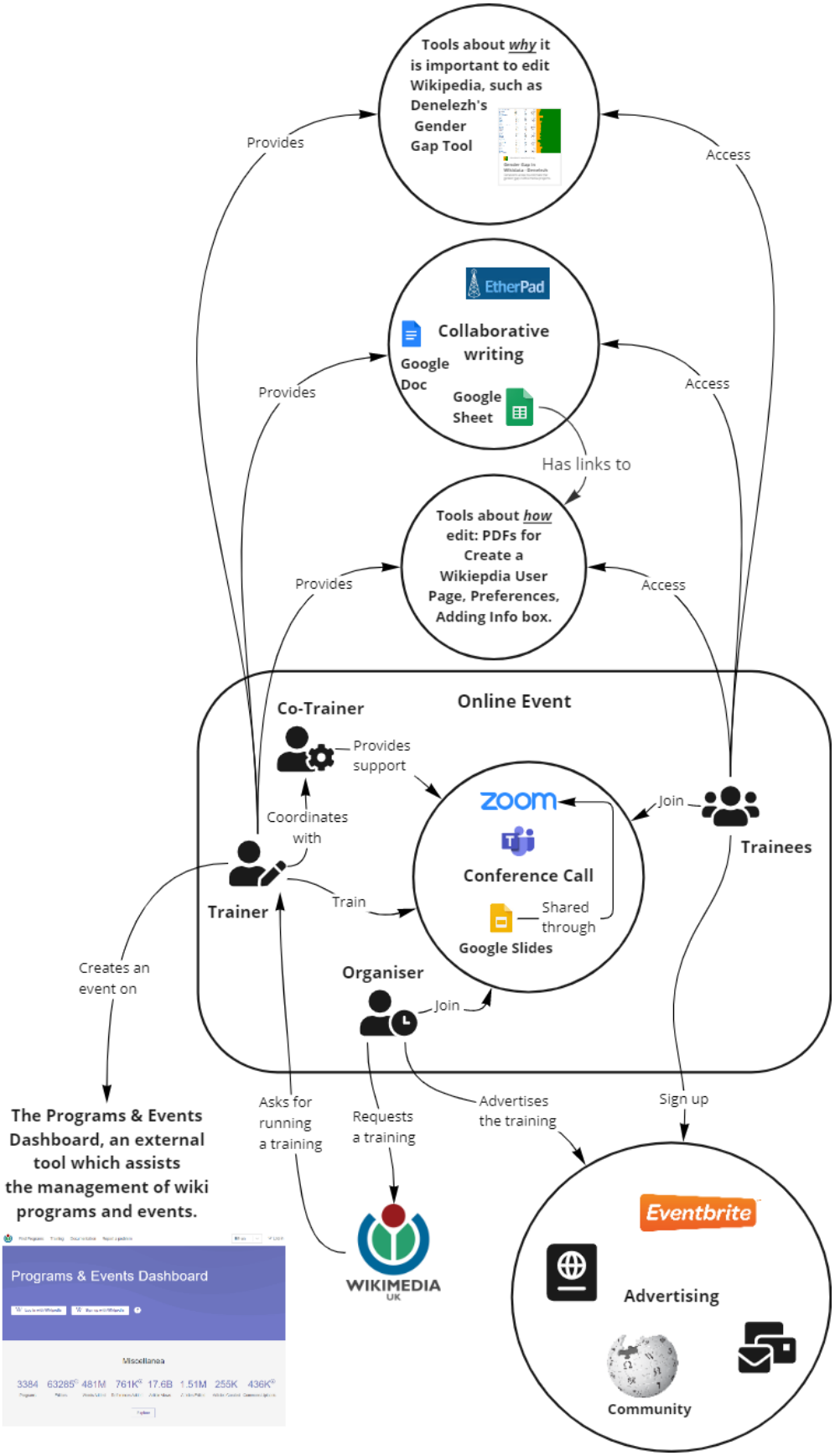


Figure 4.18 Flow diagram of a virtual training and the different systems used by trainers, co-trainers, organisers, and trainees.

This shows the diverse needs of trainers and above-average technical expectations of the trainees. This also helps answer the first question by designing the online Wikipedia training activities flow diagram, shown in Figure 4.19, which illustrates a Wikipedia training process for creating articles about notable women based on observational data. Each lane in these diagrams shows the activities performed by a specific actor. Trainees are at the top in teal-ish blue, then the Trainer/s lane in yellow, followed by the Organiser/s in red, and lastly, the Speaker/s or Experienced Editor/s in green.

Section 5.4.3 discusses how this diagram helps form the basis of an improved diagram shared with the participants in the sessions, playing a key role in co-designing a scenario for introducing real-time collaboration.

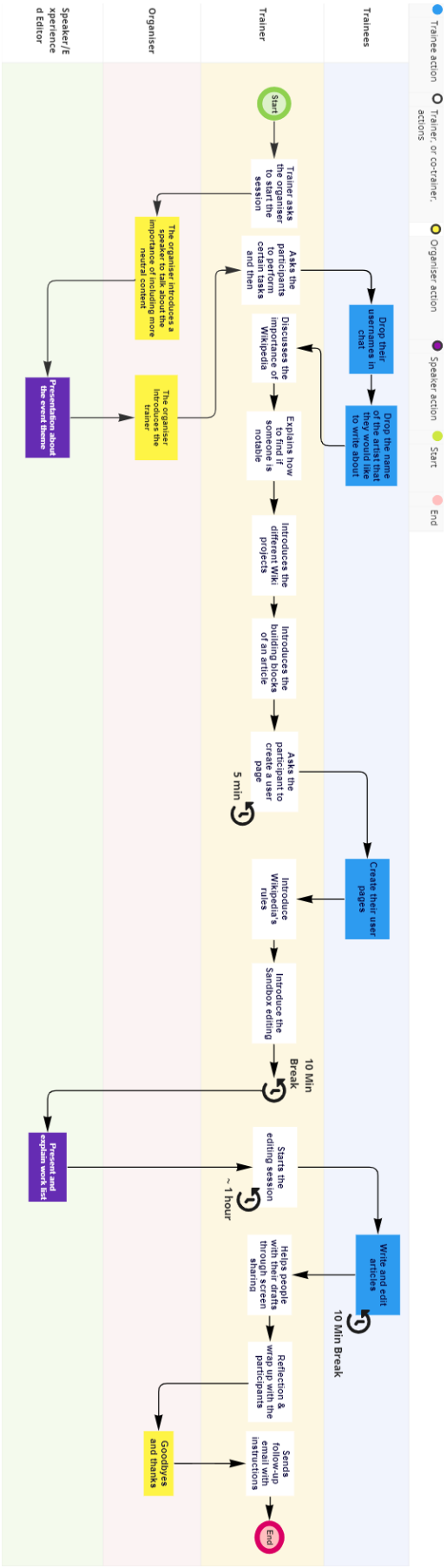


Figure 4.19 An online Wikipedia training activities flow diagram.

Q1.5 What are the collaborative writing modes during training?

Based on observations, writing a Wikipedia article in online training could be divided into three stages. The first starts before or during the training, where the organisers encourage real-time coordination on finding articles to write and resources to use. However, in the second stage, participants are asked to write the articles separately to avoid editing conflict. The last identified stage is reviewing the article. This involves the trainer or an experienced editor reviewing the article in the Sandbox space to enhance or "Wikify" (adding citations, categories, links) parts of it or even leading the participant verbally in a step-by-step process to move it to the main Wikipedia. This might involve sharing links in real-time chat or asking the attendees to share their screens as an easier way to guide them through the steps.

The only alternative model of collaborating on the article would be when two trainees work on different parts in the second stage to be merged by one trainee for publishing later.

Q1.6 What are the challenges and opportunities for newcomers and trainers?

Many challenges have been discussed in learning how to become a Wikipedian without receiving training. While training helps one avoid uncertainty and confusion, it poses its own challenges, too. Most such difficulties were communicated directly using the key observations section. Below is a summary of some of them.

For Trainers:

- After years of face-to-face training, a trainer was a bit nervous about running the event online. It was also challenging for her/him to cover everything needed to develop someone into a fully-fledged Wikipedian in one 3 to 4-hour online training session.
- A trainer must keep track of many elements (presenting, watching chat, moving between tabs to demonstrate something), where many things could go wrong in the training process.
- Technical know-how can vary among the attendees, and it is difficult to design flexible training to fit these needs.
- A trainer cannot pick up on the non-verbal cues that could let her/him know that a trainee is lagging behind.

For Trainees:

- Switching between Wikitext and the Visual Editor is one of the most confusing elements of the editing process for attendees.

- Wikipedia is for everyone and can be edited by anyone, however, the user segment, which can take a 3 to 4-hour block of time out of their schedule to learn how to edit Wikipedia, is limited.
- Trainees who see themselves as fitting into the community's stereotypical image of a Wikipedia editor are above average in technical skills and sometimes "geeky". "The type of people who were drawn to writing an encyclopaedia for fun tend to be pretty smart people", according to Jimmy Wales (2007).
- It is challenging to find the value of investing hours in learning how to edit Wikipedia, which has been the point of criticism by academics for years.
- Occasionally, trainees are subject to assumptions of being able to use specific features, such as screen sharing. Asking how to perform such tasks in front of others in an online conference call could be daunting or embarrassing. Another challenge is being unaware of one's need to speak through her/his body language or facial expression. Similarly, it is challenging for a trainer to pay attention to non-verbal cues when someone is stuck.
- However, many Wikipedia training opportunities were observed, including:
    - The use of the real-time chat functionality afforded many possibilities to support collaboration between trainer and trainees, as well as among trainees.
    - In one of the training sessions, a trainer shared a link to Google Slides that allowed the participants to follow the trainer's progress in real-time and access the links displayed in the presentation.
    - An early collaboration on Google Sheets that has the articles worklist led to more coordination and heightened the attendees' commitment to join and work on the articles, who shared their findings with the rest of the group after more research on the worklist.

4.5 Implications for the Participatory Design Process

One of the PD challenges that have been identified "is that of developing complementary means of 'taking a closer look'." (Crabtree, 1998, p. 3). Crabtree argues that relying on user participation is insufficient to find transformational technology solutions and may lead to "finding the perfect solution to the wrong problem". According to Crabtree, the solution could be employing ethnography to display the details of "what the work is really all about".

4.5.1 Ethnography's Contribution to Scoping This Research's PD Process

Gaining more contextual knowledge about Wikipedia training has helped scope the study and identify different communication channels with stakeholders.

The participants worked individually, with some having difficulties using features such as screen sharing; therefore, the early sessions should examine the technical abilities required to go through a real-time collaboration process. This would need broad trainer participation and input, as the observation session was limited to staff members at Wikimedia UK. It would also open an interesting discussion into how having different roles based on the trainees' backgrounds would help the group collaborate on an article. For example, one participant might learn how the Wikieditor/Wikitext works, while the others would take different tasks that do not involve dealing with Wikieditor.

4.5.2 Why is Real-Time Collaboration Missing From Training Events?

This observational study and the initial Activity Flow (see Figure 4.19) have provided an opportunity to ask a more detailed and on-point question using correct community language and terms in the sessions. As will be discussed in Chapter 5, the Activity Flow diagram serves as a conversation starter, and the PD participants complemented it with several missing activities focusing on editing and writing articles. A clear flow of what regular training resembles has helped centre the discussions around activities that could be affected or done differently had real-time group collaboration been introduced. It has also offered the necessary partial view of the training to see how the early planning phase affects the collaboration pattern during training.

4.5.3 Avoiding the Technical Challenges and Building Collaborative Groups

Identifying the various technical backgrounds of participants in the observational study shows that assumptions about what the newcomers know could isolate those below this assumed standard. Therefore, the ideal situation for a trainer to conduct successful training is to have a group of participants with similar or close technical know-how levels.

To introduce a real-time collaboration model in Wikipedia training, one needs to experiment with creating and identifying one setting or scenario that real-time could support participants and later expand to explore other cases. Therefore, the sessions of Chapters 5 and

6 explored with participants how to decide on a group of participants with similar or close background knowledge that could benefit from real-time collaboration in a real-life scenario.

4.5.4 Revisiting Assumptions

At every observed training session, the trainer asked the participants not to work on the same article synchronously to avoid conflict of edits. I would argue that explicitly setting the norm that synchronous collaborative writing is an undesirable act would lead to edit conflicts and decrease the chance of newcomers exploring this option in the future through the Wikimedia innovation process. This was explored further with newcomers in the sessions.

4.5.5 The Solution Could be a Middle Layer Between Wikipedia and The Volunteers

Collaboration on resources demonstrates the opportunities that real-time communication could bring. Many of these features have been discussed in the community but did not proceed due to social and technical complexities. Therefore, the PD process could explore with participants situating the solution in a platform that would function as a middle layer between Wikipedia and potential volunteers, creating volunteer training. This led to the conceptual solution model covered in Chapter 7.

4.5.6 Distributed Participatory Design

Distributed PD has proven an effective tool for communities to collaborate on designing solutions beyond the COVID-19 challenges of bringing participants into the same physical spaces (Slingerland et al., 2022). However, the question of how many participants should be involved to have a global, inclusive design cannot be more complex with a community spanning over 300+ languages. Moreover, using traditional PD techniques to involve the Wikipedia community would be impractical, and finding solutions to a large project should not rely merely on conventional PD practices (Oostveen & van den Besselaar, 2004). Therefore, this research must use a combination of established PD tools, such as workshops and scenario-based evaluation with large-scale social voting and ideating on these workshop results to include the broader voices of the Wikipedia community to avoid "designing the perfect solution to the wrong problem." (Crabtree, 1998, p. 10).

4.6 Summary

The outcome of this chapter shows the importance of understanding the context better through exploring the software development process and observing Wikipedia training.

Discussing MediaWiki's innovation process demonstrates Wikimedia's design process and its capacity to involve the community in it. In addition, it shows the high-level technical know-how needed to participate in the process. The Wikipedia community's global presence and diverse views also add another complexity. This has led to addressing why real-time collaboration feature requests did not build the momentum to reach the end of Wikimedia's innovation process funnel.

On the other hand, this chapter solidifies the case for benefiting from the opportunities Wikipedia training provides as a tool for introducing change and preparing newcomers to be part of the Wikipedia community. Also, observing the training has helped draw a picture of its participants' challenges, tools in use, and their approaches to using real-time tools, such as Etherpad, Google Sheets, and a conference chat feature in coordinating work. In addition to the confusion of working together or alone and switching between Wikipedia's Source and Visual editing, the trainees have received clear but contradictory clues on how to collaborate on writing Wikipedia articles.

Exploring the innovation process and Wikipedia training has made it possible to design a large segment of the early sessions discussions and questions that needed to be asked. Examining the community's proposals in the Wishlist and strategy or Wikimedia's IT specialists in the Phabricator, one can see that each group has a unique view on implementing the real-time collaboration feature in Wikipedia. Undergoing the sessions equipped with all of these different views has helped start the discussion where the debate about real-time collaboration had stopped. Observing the training session also contributed to the sessions, leading to a focus on exploring possibilities and arguments to revisit assumptions about existing work practices and the collaboration model and how they could be co-evolved.

Thus, the observation study described in this chapter is used in shaping the design of the sessions in Chapters 5 to 7 for designing a tool that involves real-time collaborative writing.

Chapter 5 Discovery Phase

The ***Observational Study*** of Wikipedia training sessions has produced a vivid picture of the context of this research and explores community challenges, collaboration models and workarounds using third-party real-time collaboration tools as well as piecing together the design practices for Wikipedia software. After an expert review of the outcome, the research project was ready to move to the next phase. As shown in Figure 5.1, this chapter covers the ***Discovery*** phase, discussing the "Identify" and "Define" sessions and techniques that have helped discover the challenges and potentials in Wikipedia training collaboration models from the community perspective.

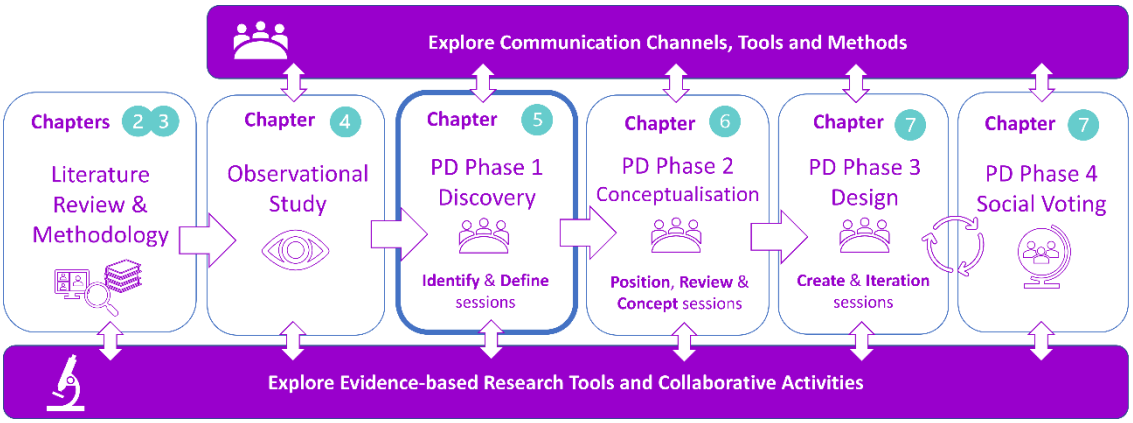


Figure 5.1 The Discovery phase is highlighted in relation to other stages of this research study.

Even though the research focuses on how to introduce real-time collaboration to Wikipedia, these early-stage sessions of identifying the challenges of Wikipedia training context were not particularly concerned with guiding the users directly to find the answer to **Q2 What are the key features of a real-time collaborative tool for Wikipedia newcomers?**

Q2 is divided into the following list of sub-questions answered throughout this chapter's sessions:

Q2.1. What do trainers and active Wikipedia editors perceive as challenging with Wikipedia's current collaboration model and tools?

Q2.2. What are the challenges and the context of possible solutions with introducing real-time collaborative writing in Wikipedia articles from the community perspective?

Q2.3. How to design a study that would help explore the need for and the context in which real-time collaboration would be used?

Addressing these sub-questions establishes the base that the following chapters will rely on, leading to defining the primary design decisions in 6.7.1 and the features of a real-time collaborative training tool prototype covered throughout Chapter 7.

This Discovery phase has resulted in discussions about other critical challenges faced by the community. Starting with the broader picture of challenges has helped connect the issues facing the implementation of real-time collaboration from multiple perspectives. In later phases, the sessions bring the focus back to the main research question.

As discussed in Chapter 3, this research benefits from Hagen et al. (2012) PD methodology in shaping a new ethnographically-informed distributed participatory design framework. In this chapter, the design of the Discovery phase is covered (Figure 5.2, which encompasses two sessions, "Identify" and "Define"). The bottom of the figure shows the activities used according to the session number.

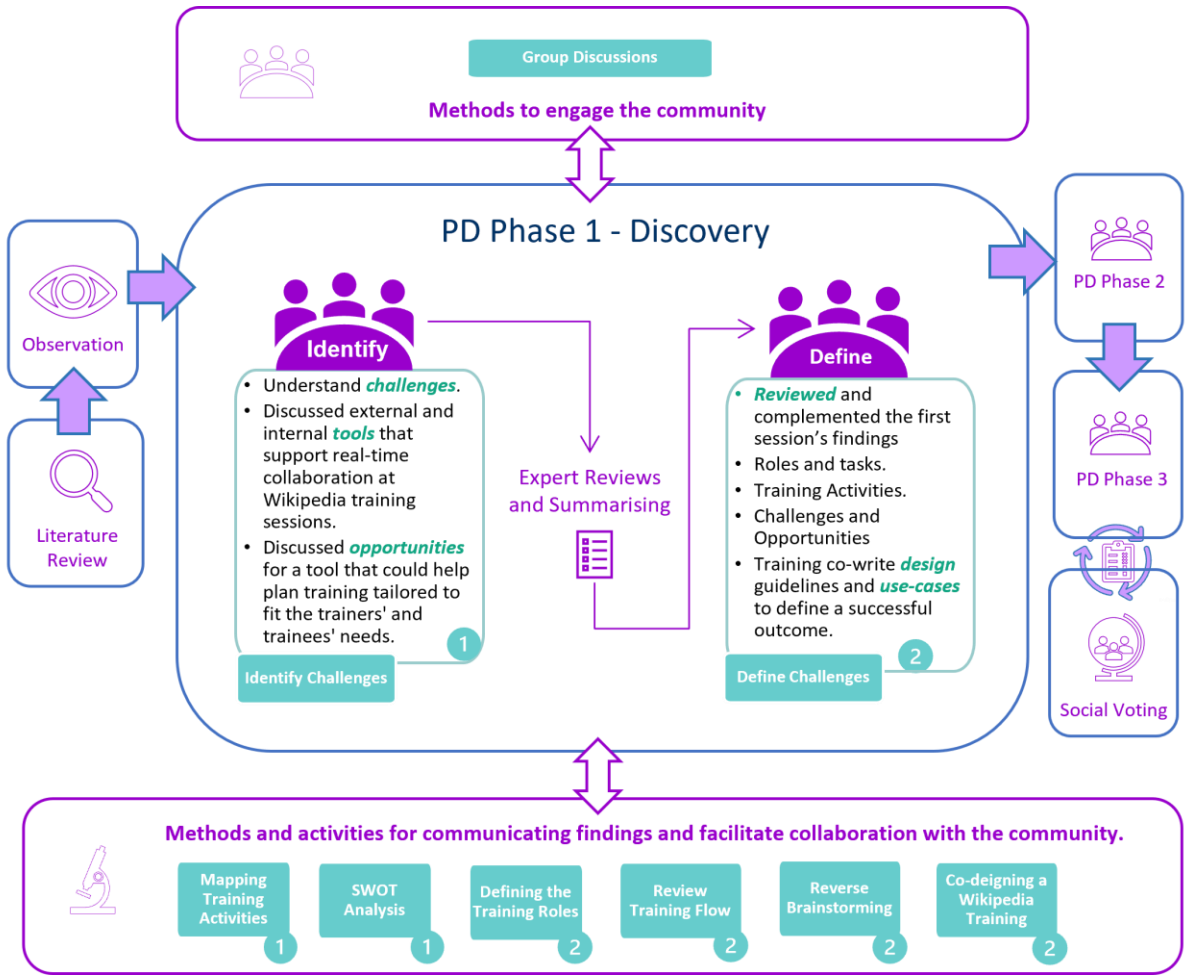


Figure 5.2 The Discovery phase and its methods.

First, Section 5.1 provides details on participants' backgrounds, data collection approach, and the software used in the sessions. Next, Section 5.2 details the activities used in the first two

sessions that constitute the first Discovery phase. This leads to the analytical approach in Section 5.3, followed by the insights and findings in Sections 5.4 and 5.5, which cover each session's findings separately. Finally, the chapter concludes in Section 5.6 with a discussion about the opportunities identified, reflecting on the lessons learned and the aims of the next chapter.

Some details, such as the detailed description of the sessions participants' backgrounds, onboarding process, and the PD research software, are only mentioned once in this chapter to avoid repetition in the next two.

5.1 Participants and Data Collection

This section discusses the choices in sampling the participants for all sessions and the data collection methods, focusing on the Discovery phase. This is followed by the analysis approach used to define the insights and findings at this phase.

"Notions of who should and should not be involved are always preliminary" (Hartswood et al., 2008, p. 61). Finding participants for my research underwent several stages. The early phase benefited from my previous involvement in the community, which led to my collaboration with Wikimedia UK, a key player in training Wikipedia newcomers. Having examined the observational study findings, I identified the targeted audience for the sessions before developing the advertising material for recruitment. The Conceptualisation and Design phases had participants from several countries and different levels of experience.

The data collection plan involved the formation of a participant information sheet and an electronic form to gather background information about her/him, followed by a consent form (see Appendix D.1). After reading the advertisement (see Appendix D.2), the participants filled out the application form, which has the PIS. Having checked their information and skills, I sent an email containing a poll to arrange a suitable time (see Appendix D.3) for the sessions. Later, they received a test Miro board with instructions on navigating its interface. After joining, the participants were briefed about the research background and tools used, given the opportunity to ask questions, and encouraged to interact with the boards while joining through the conferencing tools for two hours each session.

Since Wikipedia training was identified as the context in which an opportunity exists for an alternative real-time collaboration model (see 4.1.1), it was essential to recruit potential end-users who would use the system. Recruiting non-users for this stage was excluded due to the focus of this stage on exploring the challenges and opportunities which required minimal experience with the current training model and potential users for the new training model. As

will be discussed in 8.2.1, to ensure an inclusive process, at a later stage, participation was open to non-users. However, there was no evidence of their engagement. The section also covers how recruitment for this research was one of the main challenges that led to several session cancellations, especially with newly trained Wikipedians. With the help of the research collaborators, two main categories of participants were identified: Wikipedia trainers (experienced or active editors) who have executed approximately five edits per month for the past six months, and newly trained Wikipedians who joined after a Wikipedia training, or an Edit-a-Thon, during the last three months.

This study was advertised (see Appendices D.2 and Appendix BE.2 for the recruitment note) through Wikipedia mailing lists within the UK local Wikipedia editing and collaboration communities, in addition to professional and science-oriented mailing lists.

Background information about participants' demographics is shown in Table 5.1. The letter “S” refers to the PD group session number, while the letter “R” refers to the one-to-one review sessions, and “1” or “0” indicates a participant’s attendance record. The fifth session had three iterations, so the table shows S 5-1, S 5-2, and S 5-3.

- Twenty-five participants filled out the consent form, of whom 13 were females, 11 were males, and one preferred not to answer the gender question.
- Thirteen participants indicated residing in the UK, ten were from Nigeria, one from Canada, and one from Jordan.
- Twenty-one participants had over one year of experience editing Wikipedia.
- Eleven received face-to-face training, whereas nine had received online training sessions over a year ago.
- Six have administrative experience, and eleven work at or have had a role in the Wikimedia Foundation or its chapters.
- Throughout the phases covered from chapters 5 to 7, I conducted nine online sessions that involved 15 participants, of whom 14 were Wikipedia trainers or experienced editors, and 14 participants who took part in the study had previously attended Wikipedia training. The second review session (R2) involved a newly trained Wikipedian, which is why years of experience show a null value. The majority participated face-to-face, and a few online.
- The nine sessions included two one-to-one review discussions with an experienced editor and a newly trained Wikipedian. The fifth session had three iterations. The last session introduced the final version of the WikiSync prototype, as will be discussed in Chapter 7.

• Through the research, I received feedback from three research collaborators, who also function as Wikipedia trainers or Wikipedia UK staff members, on the research direction and design choices.

Table 5.1 The demographics of the session participants according to the phases.

				Phase Title								
				Discovery		Conceptualisation				Design		
Attended a Training?	Gender	Country	Years of Editing	S 1	S 2	S 3	R 1	R 2	S 4	S 5-1	S 5-2	S 5-3
Yes	Female	England	2	1	1	1	0	0	1	1	0	0
		Nigeria	1	0	0	0	0	0	1	0	0	1
			3	0	0	0	0	0	1	1	1	1
							1	0	1	1	1	1
		Scotland	Null	0	0	0	0	1	0	0	0	0
			1	0	1	0	0	0	0	0	0	0
			8	0	1	0	0	0	0	0	0	0
	Male	England	14	1	1	1	0	0	1	1	1	0
			18	0	1	0	0	0	0	0	0	0
				1	0	0	0	0	0	0	0	0
		Nigeria	4	0	0	0	0	0	1	0	0	0
			5	0	0	0	0	0	1	1	1	1
		Scotland	5	1	1	1	0	0	1	1	0	1
No	Male	England	13	1	1	0	0	0	0	0	0	0
		Jordan	8	0	0	1	0	0	0	0	0	0
				Dec/21 - Feb/22		Feb/22 - Apr/22				June/22 - Feb/23		
				Phase Timespan in MM/YY Format								

Sample Achieved for the Discovery Phase

As shown in Table 5.1, out of the fifteen participants in this study, eight took part in the Discovery phase. Some participants signed up for the study but never participated due to time conflicts. The Discovery phase included two online group discussion sessions between December 2021 and February 2022. While the timing and date for the "Identify" session (S1) were advertised based on an initial timing poll, the "Define" session's timing (S2) and date were based on the end of the session discussion.

The "Identify" Session Participants (S1)

Out of seven participants who signed up for the first session, five (including one female) were able to participate, and two opted out an hour after the start of the session owing to pressing circumstances. All participants were UK-based and had editing experience that ranged from 2 to 18 years. Four participants had over five years of experience, and the fifth, a female editor, had two. Three attended face-to-face training, and two learnt on their own. Four of them indicated their involvement in Wikipedia training activities. One participant held a high

administration role in the English Wikipedia community, and the other four, at the time of the study, hold or previously held positions in the Wikimedia Foundation or one of its chapters.

The first "Identify" session was planned to run for 2 hours, although some time was lost at the start due to difficulties a participant had with his browser, resulting in extending the session for seven minutes. This session led to learning many lessons on designing the next "Define" session of the Discovery phase.

After the 2021 holiday season, the first session summary was shared, and invitations were sent for the "Define" session in February. It is noteworthy that most participants edit Wikipedia aside from their full-time jobs, preferring to have the sessions in the early evenings and on weekdays.

The "Define" Session Participants (S2)

Unable to run the session due to last-minute cancellations, I successfully rearranged it by recruiting nine new participants, seven of whom attended, including one research collaborator from Wikimedia UK, who helped get more participants to join to prevent any other session cancellations that might disrupt the study flow.

Data Collection

The first two sessions of the Discovery phase, which varied in length, were recorded and transcribed using Microsoft Teams and later manual refinements. This provided 4.3 hours of video data. In each session, the participants were briefed and given guidance, followed by a list of open-ended questions as part of the activities discussed in Section 5.2. After each session, Miro board screenshots were taken and uploaded, along with the videos and transcripts to NVivo software for analysis.

5.1.1 Software Used

This section discusses the online software used for sessions and how methods have been adapted to fit the online research of a distributed community.

To find the right approach for conducting the sessions, the following section discusses tools used in similar distributed PD research, leading to the inclusion of the tools and methods that fit my study.

Existing approaches and consideration

The growing interest and uniqueness of PD have engendered questioning about the availability of software to support its specific practices. Examining related research work would help define the tool/s to be used in conducting my research project.

The traditional PD methods mostly support face-to-face activities. This fact, combined with the absence of tools specifically designed to be used in DPD, has led some researchers, such as Walsh et al. (2012) and Heintz et al. (2014), to develop software solutions that simulate some online PD methods. Heintz et al. categorise the functional and usability requirements for such tools as follows:

User requirements:

- Interactivity: the ability to interact with the prototype in contrast to a static image.
- Annotation: the ability to comment on specific design points in text format.
- Creativity: providing features, such as drawing or manipulating an interface component with no limitation to their imagination.
- Collaboration: the ability to work, edit and annotate an artefact with more than one participant.
- Access: The ability to access it easily via the Internet at any location.
- Instructions: the ability to provide the participant with clear instructions on exploring the environment to get started, even without the researcher's involvement.
- The Developer Requirement
- Activity: The ability to collect the data generated from user activities (navigation patterns and mouse clicks), as well as feedback.
- Aggregation of data: supporting both implicit and explicit aggregation of data.
- Export: being able to extract the data from the tool and export it to other statistical software.

(M. M. Heintz, 2017) discuss the prototype of a tool specifically designed for a distributed PD, titled PDot. The author describes it as "a user-friendly and customised tool that can elicit similar results as paper-based approaches." (M. M. Heintz, 2017, p. 132). Among the other features influenced by user requirements, PDot, as shown in Figure 5.3, offers an interface with annotations feature that allows end users to comment on the designs. This has brought on researching available tools with similar characteristics to facilitate my distributed PD sessions.

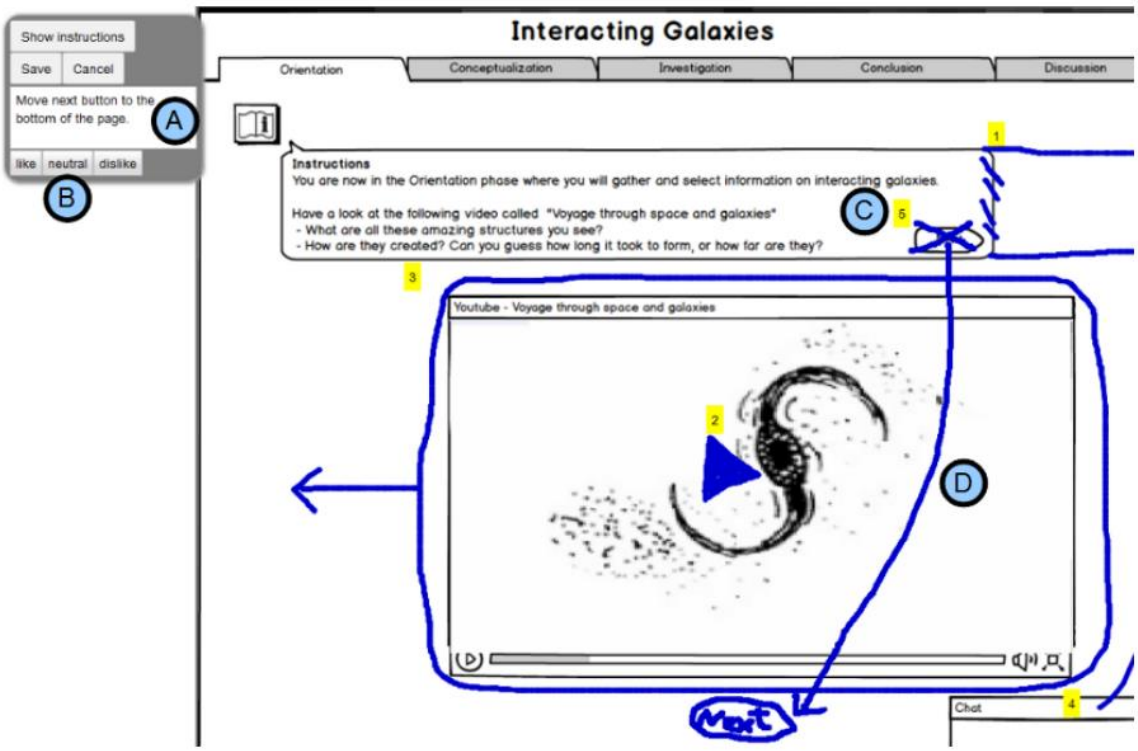


Figure 5.3 PDot tool interface that allows users' annotation (M. M. Heintz, 2017).

Miro

The COVID-19 pandemic has led many PD-oriented research projects to adapt brainstorming and PD practices to a distributed online research context, benefiting from the flexibility of online services, such as Miro. After researching other online collaborative whiteboards, such as MURAL[72] and Figma,[73] and comparing their feature to Miro, it was found that the former Miro would be a viable option to proceed with organising this paper's sessions.

Among other useful features, Miro covers most of the usability requirements expected from a PD tool. The first, second and fourth requirements from (M. Heintz et al.'s, 2014) recommendations are covered in Miro, which allows a group of guest users of this virtual board to collaborate on interacting, moving, copying, adjusting the size of prototype elements, and commenting. Miro allows users to move through different parts of the board, and when the "bring everyone to me" feature is clicked, everyone's attention is drawn to the section discussed, increasing collaboration efficiency. The commenting could be done by typing text or drawing, which does not limit the guests' creativity, as discussed in the third requirement by Heintz et al. (2014).

[72] https://www.mural.co/
[73] https://www.figma.com/

Miro covers the fifth requirement of being accessible online, though some Wikipedia community members are protective of their identities, especially when using non-opensource software. Miro's advantage is providing co-design sessions with some flexibility in that participants are able to access the board without the need to create accounts or reveal their identities.

Regarding the seventh requirement, Miro offers a tour that guides new users through virtual space. For this reason, a space was added to the board landing area with additional instructions on its use.

Thus, using Miro in this research has proven to efficiently address each session's needs.

Pen & Paper

Participants were informed that they could use pen and paper to describe ideas they could not illustrate on Miro. This was beneficial, especially for those who joined through the phone or whose Internet connection was poor.

Microsoft Teams

The sessions were hosted using Microsoft Teams, which has the most required functionalities, such as chat, video calling, screen sharing, transcribing, and recording.

5.2 Discovery Phase– Session Activities

The Discovery phase consisted of 2 sessions, and eight participants went through several activities based on the methods shown at the bottom of Figure 5.2. The session design and corresponding activities are covered in detail in this section, while Section 5.3 discusses the approach for analysing the data resulting from these activities.

The methodological path discussed in Chapters 3 and 4 guides the selection of the sessions' methods and tools needed to capture the scope and depth of the insights originating from this PD procedure and conducted activities.

The following sections describe the use of online sessions and the techniques to facilitate a virtual environment in the first two sessions, "Identify" and "Define". This environment allowed participants to collaborate on identifying the challenges and potentials of collaboration models in training Wikipedia newcomers, helping define new alternative models to address the challenges and exploit the opportunities.

Sections 5.2.1 and 5.2.2 below provide details on the pre-session information and activities ahead of the start of most sessions in my research, which is covered once in this

chapter. Sections 5.2.3 and 5.2.4 dive deeper into the specifications and activities that run as part of the sessions' activities.

5.2.1 Pre-session Information and Activities

Given the participants' diverse expertise, creating an environment where a group of people become a team that partly shares the same background knowledge and goals has made the onboarding process a challenge. Krueger & Casey's (2014) work guides researchers with similar difficulties in planning and implementing group work sessions. The authors describe a method for designing "easy to say, clear, short, usually open-ended, and one-dimensional" questions with difficulty and focus that improves gradually. Therefore, the sessions' planning phase is motivated mainly by Krueger & Casey's (2014) work, which helped plan and promote successful group discussions in my research as well as gradually elevate the focus to reach the key questions about real-time collaboration.

Participants were sent a summary of the research, the focus of the session, the agenda, and the tools to be used before the sessions (see Appendix F for more details). They were also given access to a step-by-step process guide to Miro. Almost all sessions had the same pre-session preparations with some adjustments. For example, the "Define" session included a dedicated sandbox to encourage participants to test the boards.

5.2.2 In-Session Introductory Activities for the "Identify" and Define Sessions

Introduction. After joining the sessions, self-introduction, introducing the Microsoft Teams functionalities, and purpose of the session, the steps seen in the Welcome box Figure 5.4 were explained to guide through testing the Miro board together with other participants.

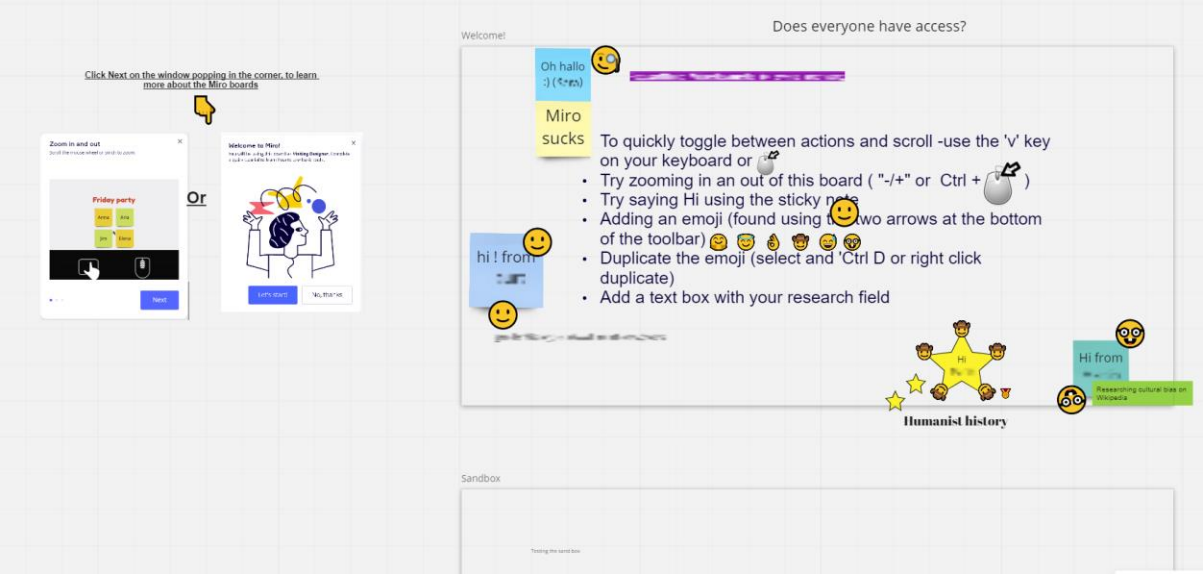


Figure 5.4 The welcome section of the Miro board used in the "Identify" session.

Name Tagging. This activity was done after the introduction to increase the awareness of the participants' individual contributions, facilitate discussions, and, ultimately, facilitate collaboration. Miro allows joining virtual whiteboards without creating an account and assigns participants a random name. Participants were asked to link their real names (or the names used to sign up for the study) to the names they were assigned as visitors by Miro, next to a specified sticky note colour which they were supposed to use throughout the session (see Figure 5.5).

Figure 5.5 Name tagging pre-session activity on Miro board.

Ice-breaker Activity. The tagging activity was followed by an ice-breaker where participants were asked to answer one or two simple questions, e.g., "What do you like the most about Wikipedia?" or "What's the best piece of advice you've ever been given about Wikipedia editing?" Participants could answer using sticky notes (see Figure 5.6 and Figure 5.7 from the "Identify" and "Define" sessions). The aim was to bring participants to communicate with each other and interact with the Miro board. The Answers were discussed using the video conferencing function in MS Teams.

This helped the participants experiment with the board, created the bond needed to start the session's research activities, as well as motivated them to start answering easy questions in accordance with (Krueger & Casey, 2014) recommendations.

The ice-breaker activity was applied in the first two sessions and dropped later as participants became used to the Miro board and knew each other by then. For the rest of the sessions, a simple round of introductions was sufficient to make new participants comfortable collaborating with others.

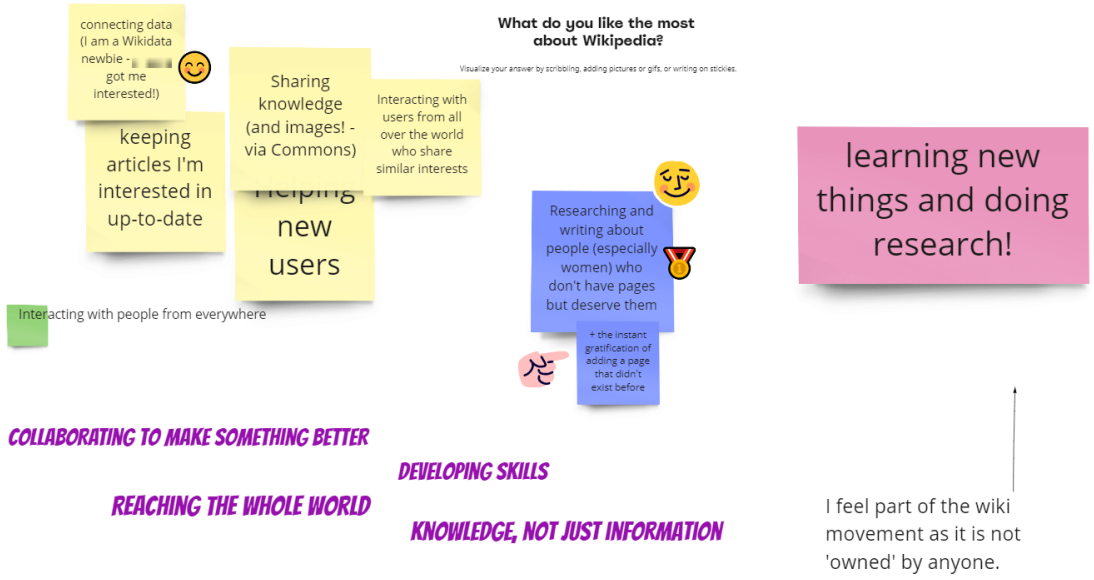


Figure 5.6 The ice-breaker activity section of the Miro board used in the "Identify" session.

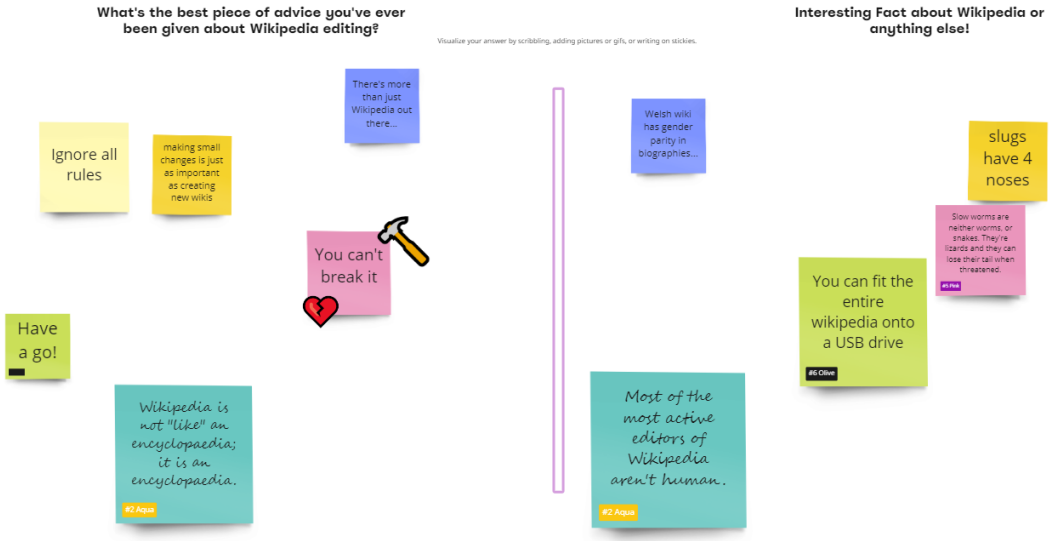


Figure 5.7 The ice-breaker activity section of the Miro board used in the "Define" session.

5.2.3 "Identify" Session – Activities

Identify. The "Identify" session aimed to identify the challenges of collaborative Wikipedia training and editing. The goal was to understand what trainers and active Wikipedia editors perceive as troublesome, given Wikipedia's limiting current collaboration model and tools.

The first session was designed to create an online environment where such challenges emerge from the end-users' perspective. The aim was to empower them to find connections to help them understand the underlying issues halting the introduction of real-time collaborative writing rather than the ones on the surface. This involved getting the participants' reflections on evidence-based research questions and findings to support their decision about the challenge this research should focus on.

The session discussed activities, such as planning for, running, and following up on a Wikipedia training session. Later, the strengths, external and internal challenges, and opportunities of online Wikipedia training were discussed, focusing on real-time technologies. This covered external and internal tools that support real-time collaboration at Wikipedia training sessions, in addition to opportunities to find a tool that would help plan tailored training that fits trainers' and trainees' needs.

Several techniques were used here that range in scope and aim. Figure 5.8 shows the full landscape of the activities at the "Identify" session's Miro board, which is covered below in detail.

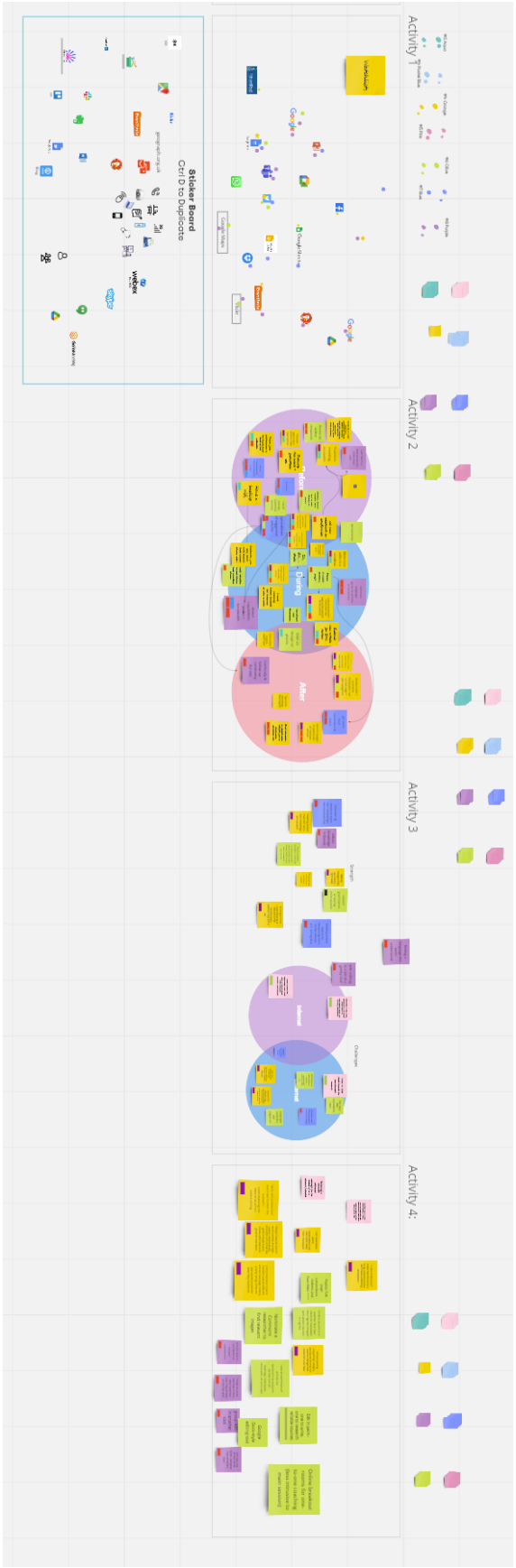


Figure 5.8 The entire Miro board for the "Identify" session to identify the challenges.

Identify: Review of the Ecology of Wikipedia Training Tools Activity

Most research on Wikipedia's collaborative writing concentrates on one tool or aspect of the system artefact ecology. Research that understands using different applications or devices to perform collaborative writing tasks is "sparse" (Larsen-Ledet et al., 2020).

In my observational study, I found that Wikipedia trainers use multiple tools to deliver training based on many contextual factors. This is supported by (Bødker et al., 2017), who argue that the artefacts' ecology is developed through time by the members' contributions and knowledge. In order to verify the findings from my observational study, in the first session, "Identify", participants were asked to review the landscape of the tools used as part of Wikipedia training sessions: what tools does the community rely on to train newcomers, whether in face-to-face or online training?

Unpacking Wikipedia training's artefact ecology and examining the use of internal and external tools of Wikipedia informs additional questions about introducing real-time collaboration features that align with the ecological development of the community's technical needs.

Drawing on work by Bødker et al. (2017) and Andreini & Forbes (2021), I involved the participants in mapping their interactions using multiple artefacts over a period of time to achieve their objectives, the Miro board was prepared, and the Sticker Board filled (see Figure 5.9) with all the tools I identified in the observation study being used with onboarding Wikipedians, more specifically, collaborative writing.

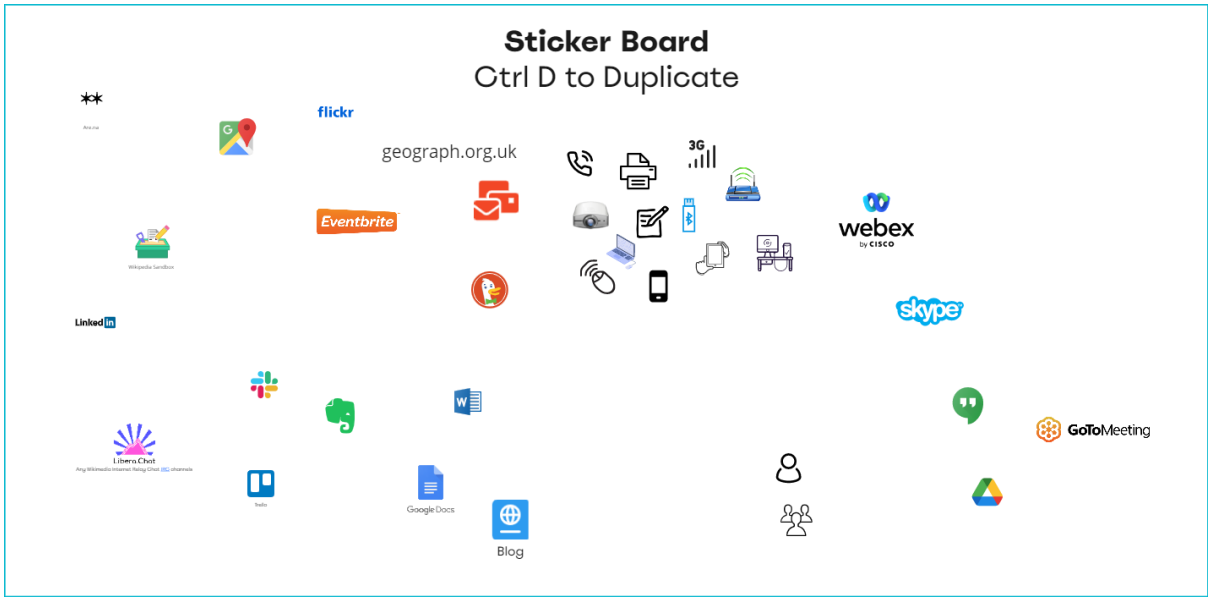


Figure 5.9 Sticker board filled with tools used in training placed on the "Identify" session's Miro board.

Tools were loosely grouped based on shared features, although these groups were not labelled. Participants were asked to move tools usually used in Wikipedia training from the "Sticker Board" space in Figure 5.9 to the "Activity 1" space on the board shown in Figure 5.10.

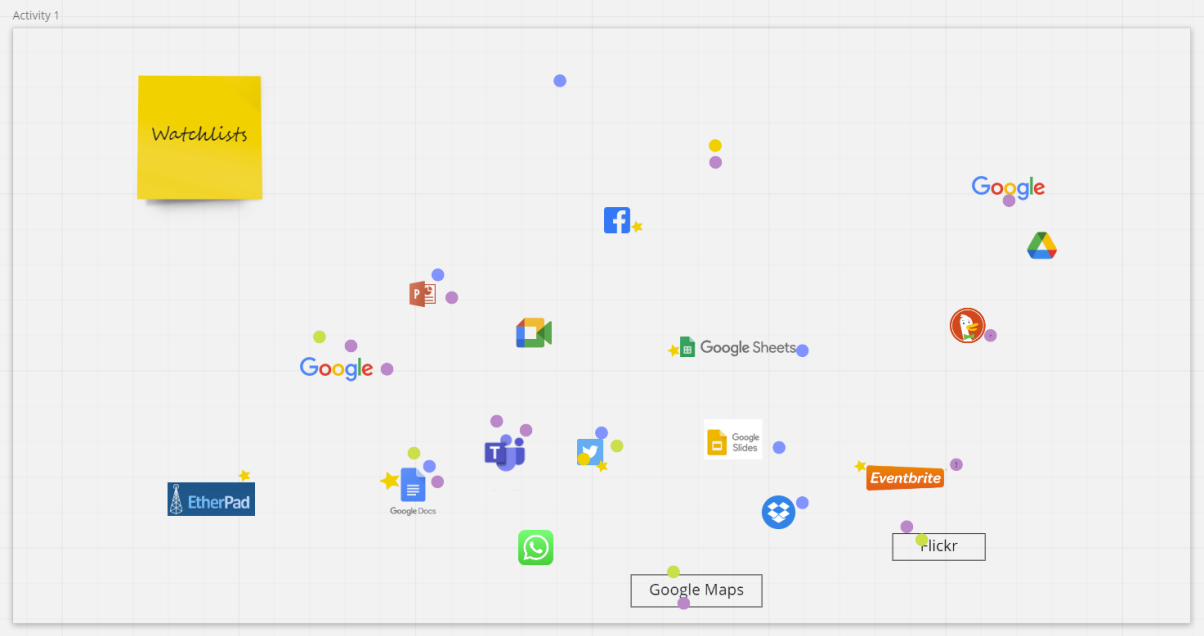


Figure 5.10 Activity 1 space to fill with stickers of the tools they use available from the stickers board of the "Identify" session's Miro board.

The participants were then asked to expand the tools presented to include those used to facilitate collaborative editing and those external to the Wikipedia tool ecology. Finally, after listing the tools, participants were asked to mark with a star those they had personal experience with as part of their collaborative writing process.

The stars on tools, such as Twitter and Google Docs, indicate that the participants have used the tool to facilitate collaboration before, during, or after Wikipedia training. Moreover, some participants added tools not found in the Sticker Board (see Figure 5.9), such as Google Maps, Flicker, and geograph.org.uk. One participant added the Watchlist[74] feature used to get notification when someone edits a Wikipedia article as part of the MediaWiki software that runs Wikipedia.

Identify: Mapping the Wikipedia Training Activities

The first activity aimed to help evoke participants' memory about the tools they used for collaboration and initiate discussions about the implicit activities or tool usages they might take for granted. The second activity of the "Identify" session focused on outlining the activities the

[74] https://en.wikipedia.org/wiki/Help:Watchlist

tools are used for before, during and after Wikipedia training sessions. It included Wikipedia's training activities, such as tailoring the training according to the trainees' technical background and setting their expectations of the training outcome. The participants were asked to reflect on training sessions they attended, performed, or merely heard about and place the sticky note about training activities in one of the intersecting circles shown in Figure 5.11.

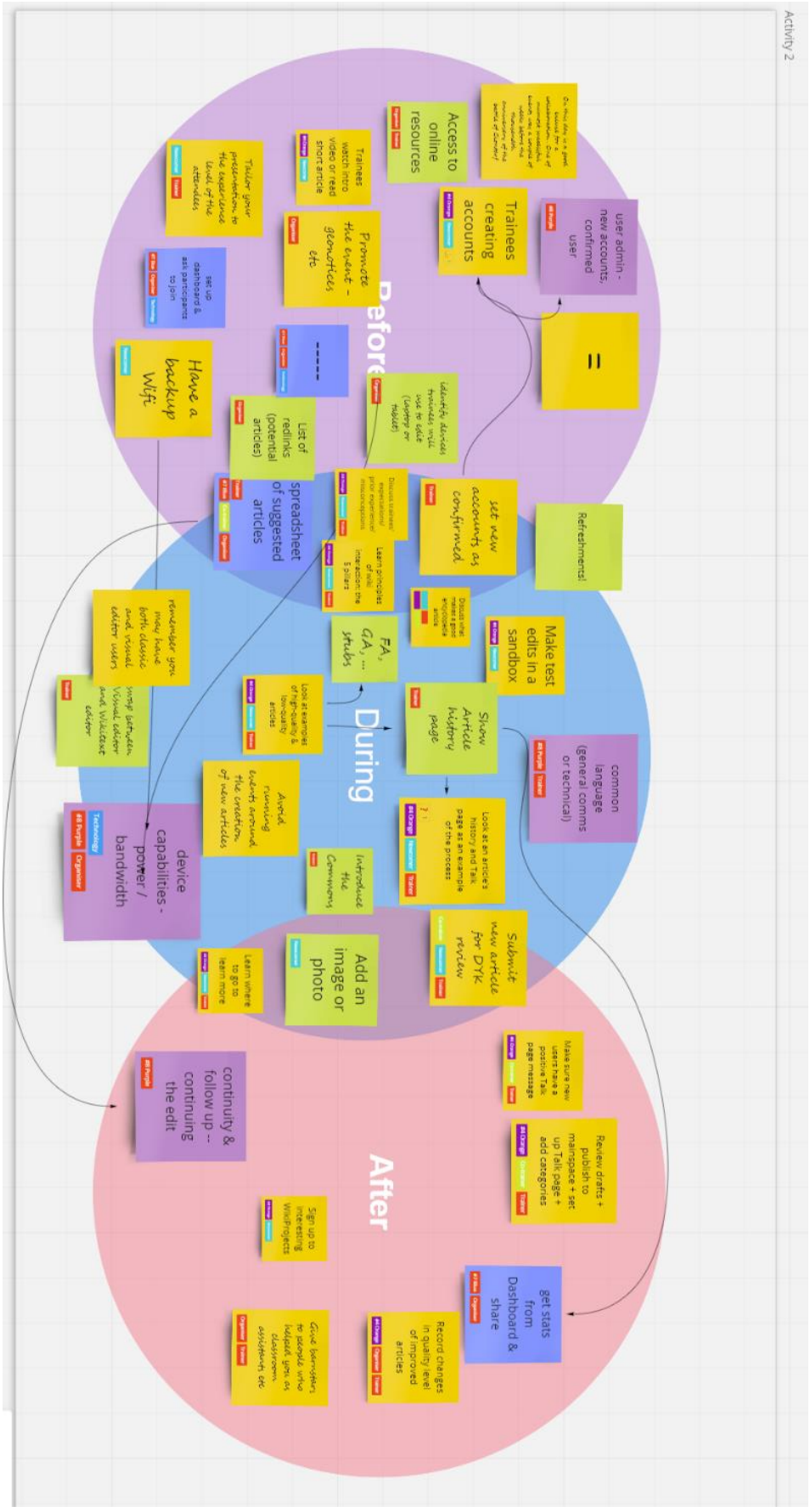


Figure 5.11 Activity space for mapping the Wikipedia training activities at the "Identify" session.

The participants were then asked to collaborate on categorising these activities according to the role of the person performing them, using the tagging feature on Miro (see Figure 5.12). Hagen et al.'s framework recommends feeding the participants' active engagement methods with an outcome of evidence-based research. Therefore, as part of the mutual learning process between the participants and me, the list of tags had some high-level themes and roles that emerged from the observational study discussed in Section 4.4, which helped participants make informed decisions. For instance, tags with the identified roles of those involved in training and other tags about the activity were included, whether technology-related or content-related. The participants were also asked to merge similar activities, bring them close to each other or link them through the thin black arrows shown in Figure 5.11.

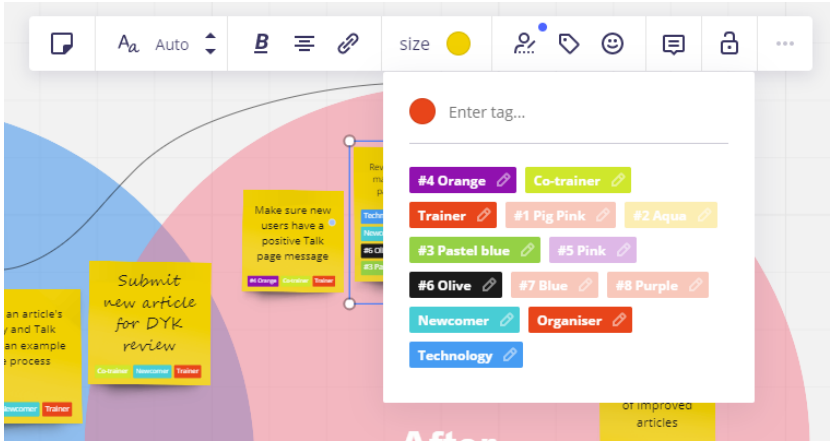


Figure 5.12 The tagging feature on Miro filled with tags.

Identify: SWOT Analysis Part 1 (Strengths, Weaknesses, and Threats) Activity

Strengths, Weaknesses, Opportunities, and Threats (SWOT) analysis is a well-established method in the strategic planning and business domain. It is used to assess the situation and create an action plan for securing a competitive position that makes use of the strengths and opportunities to address weaknesses and mitigate external threats (Leigh, 2009). This method has also been used in the research contexts, such as analysing features of the Moodle educational platform (Usov et al., 2020) and User-Centric design (Ma et al., 2007).

The session herein used a modified version of the SWOT analysis, which was split into two parts; the first was covered in the third activity of the first session, "Identity", including identifying the existing training process's strengths, weaknesses, and threats. Participants were asked to list the strengths of the current Wikipedia training flow diagram (see Figure 5.13) and discuss the unique aspects of onboarding newcomers through Wikipedia training. In addition, as shown on the right side of the figure, the SWOT analysis model's "weaknesses" and

"threats" were referred to as internal and external challenges in the session to make it easier for participants to understand its meaning and connect it back to their experience.

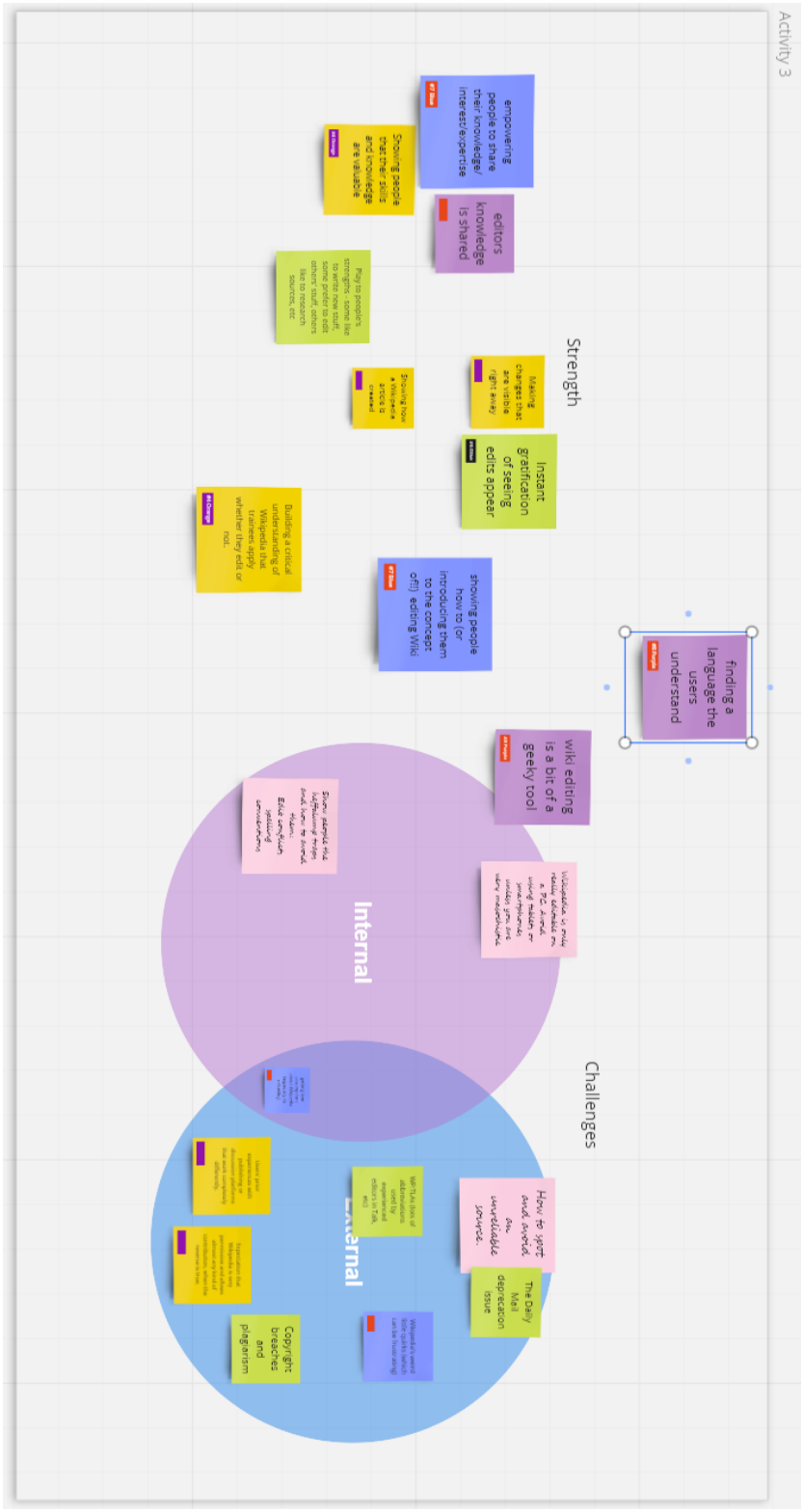


Figure 5.13 The board section shows the strength and challenges part of the SWOT activity from the "Identify" session.

Identify: SWOT Analysis Part 2 (Opportunities) Activity

In this fourth activity of the "Identify" session, we discussed opportunities separately from the earlier activity that examined the strengths, weaknesses, and threats. The second part of the SWOT analysis was the last activity of this session, focusing on exploring and sharing ideas and opportunities ranging from improving the training process to identifying the possibility of introducing a real-time collaboration model in Wikipedia. Figure 5.14 shows the part of the board where participants collaborated on the opportunities. This helped focus on the potential of having a real-time collaboration model in Wikipedia training, as discussed later in Section 5.4.4. This included discussing the pros and cons of having a synchronous editing feature in Wikipedia training events.

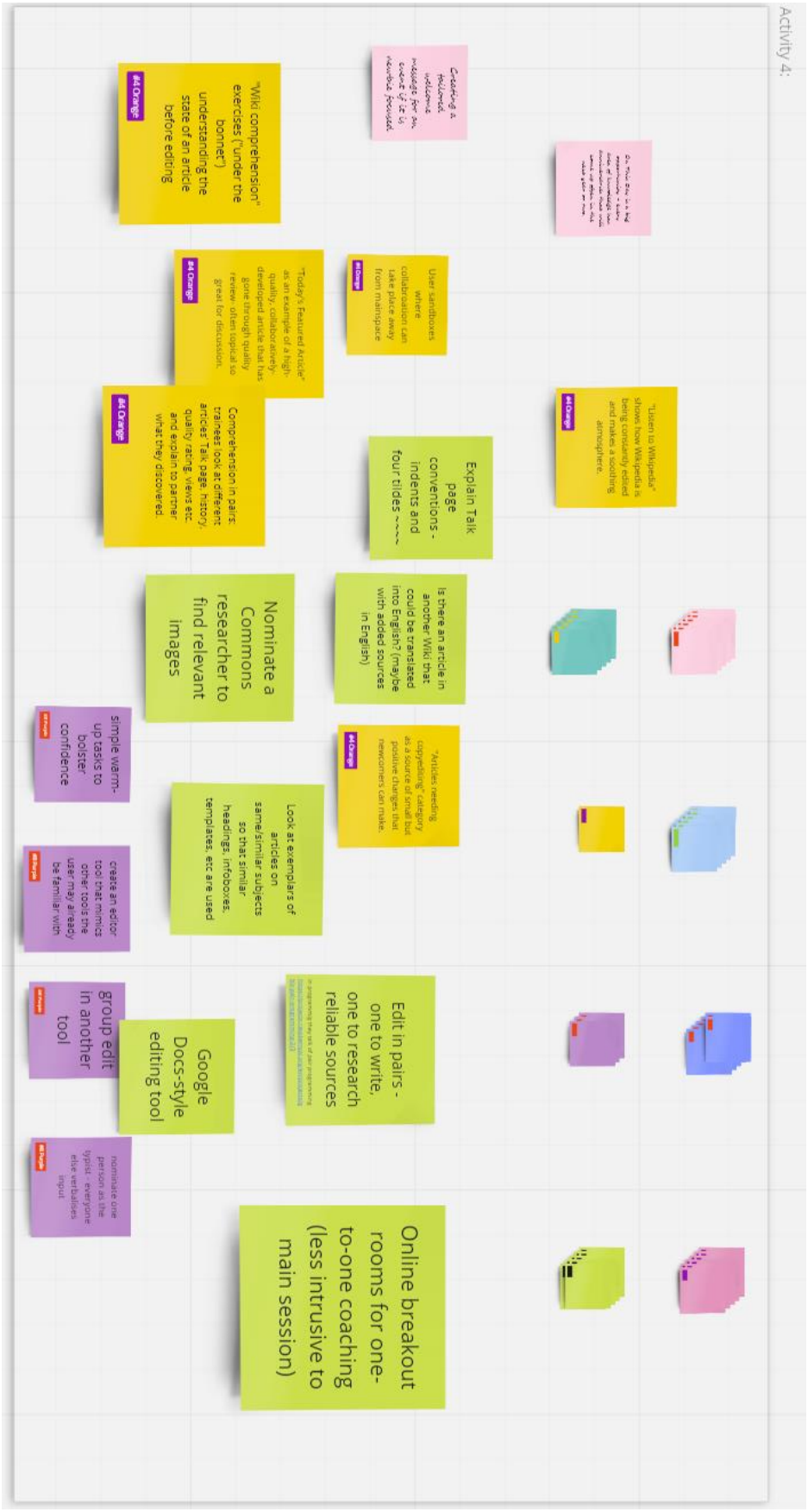


Figure 5.14 The board section shows the opportunities part of the SWOT activity from the "Identify" session.

5.2.4 “Define” Session – Activities

Define. The second session, “Define”, reviewed and complemented the first session’s findings and defined training scenarios that would benefit from the opportunities covered in the “Identify” session to reduce the training challenges. Figure 5.15 shows the full landscape Miro board of the activities, covered below in detail.

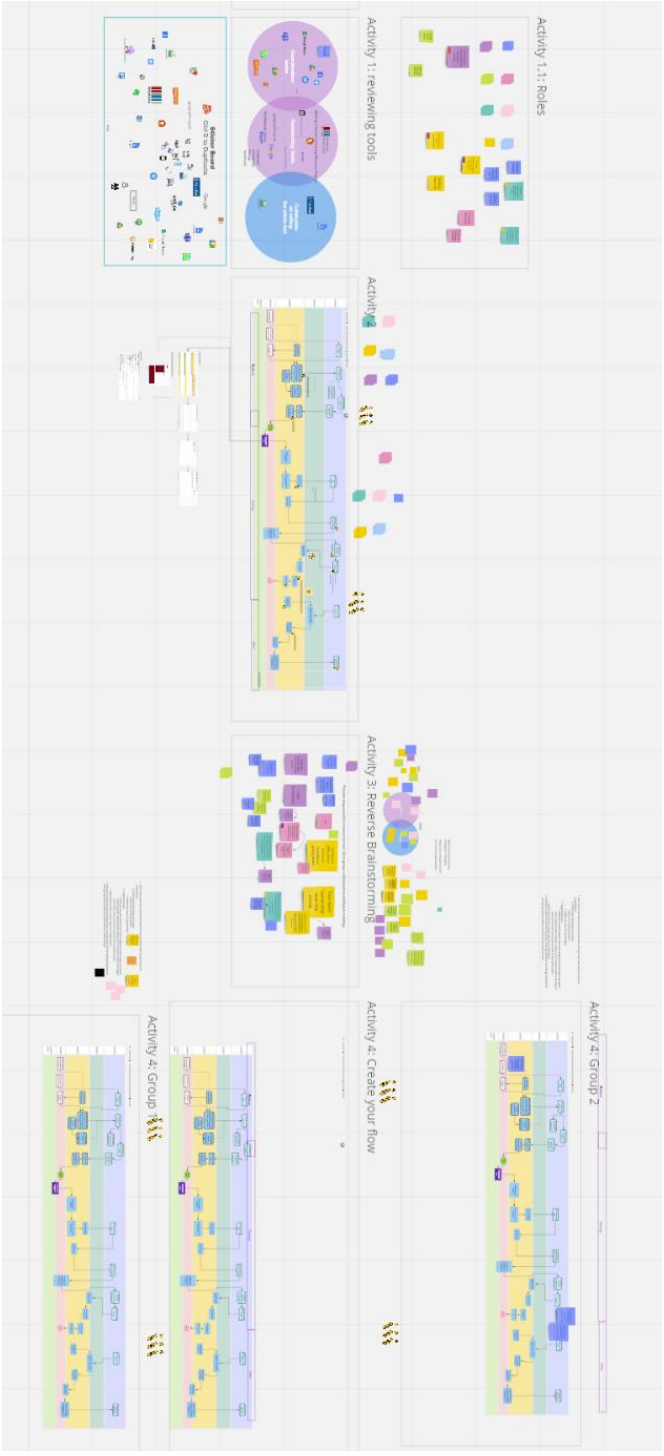


Figure 5.15 The entire Miro board for the “Define” session.

Define: Mapping Collaborative Editing Tools and Roles in Wikipedia Training

For the first activity in the "Define" session, the tools outlined in the "Identify" session (see Figure 5.10) were expanded to include tools external to Wikipedia mentioned by participants to help Wikipedians collaborate on editing Wikipedia. Then the tools were grouped into three categories, as shown in Figure 5.16: 1) tools mentioned for their role in coordinating the training, 2) the tools that support the trainees with researching content to write about, 3) and the tools used in the actual collaborative writing of the article.

The first activity of the "Define" session started with exploring any missing tools from Figure 5.16 that could facilitate the collaborative writing of Wikipedia articles.

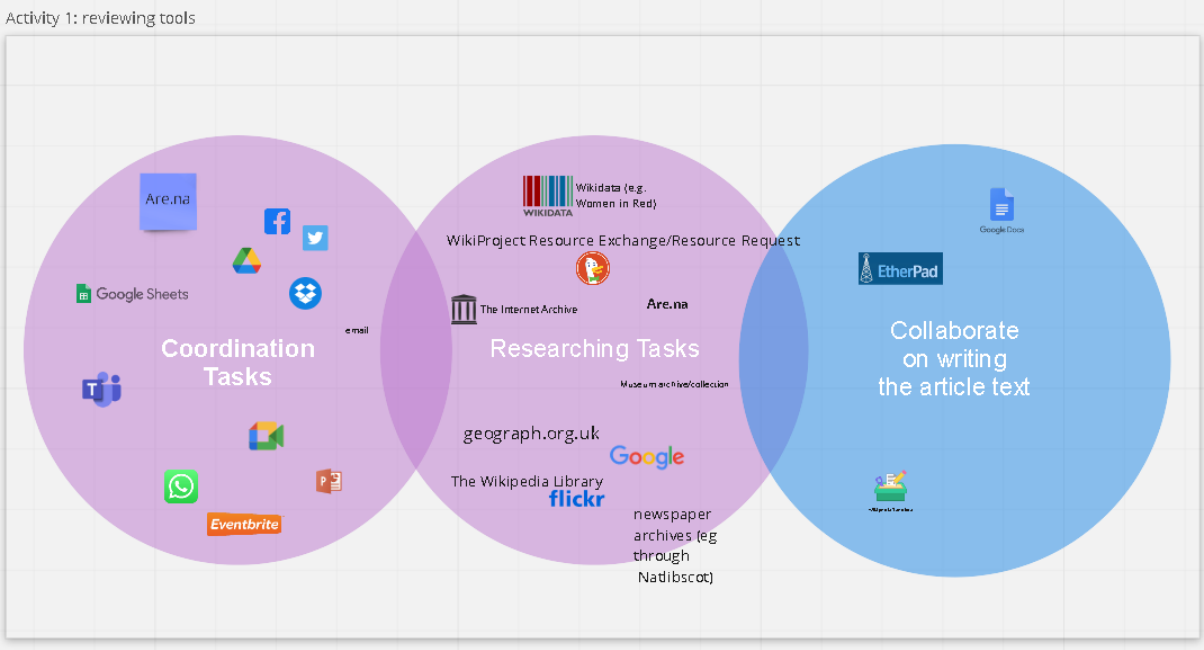


Figure 5.16 The first part of activity 1 of the Miro board in the "Define" session.

Earlier in the "Identify" session, many participants discussed assigning specific roles to train attendees to collaborate on the articles. This led me to add a second part to this activity (see Figure 5.17), where participants were asked to list roles and tasks that involved writing an article collaboratively.

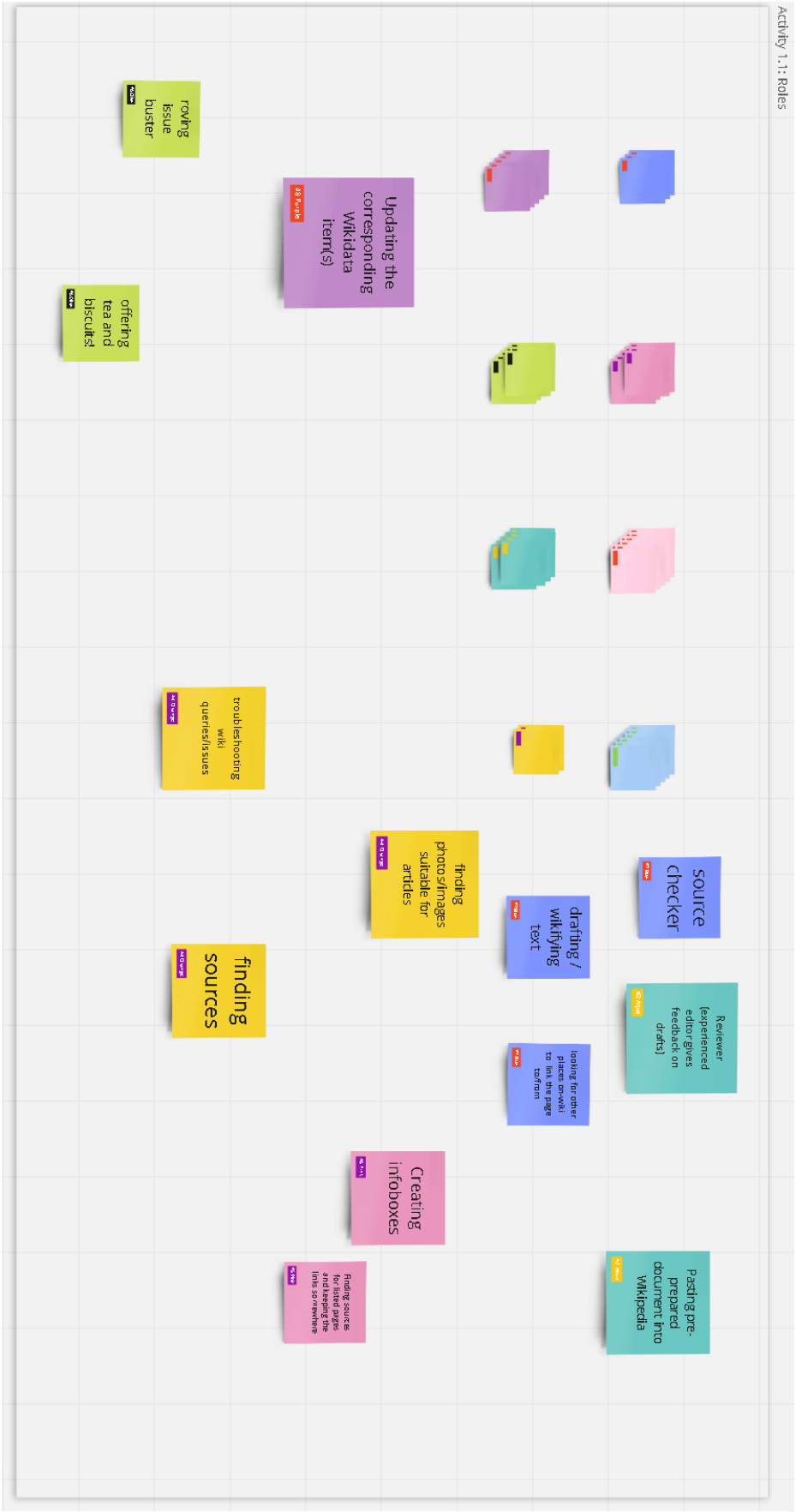


Figure 5.17 The second part of activity 1 of the Miro board in the "Define" session.

Define: Review Wikipedia Training Activities' Flow

Activities Flow Review. The findings of the first Session, "Identify", combined with those from the observational study (see Chapter 4), resulted in a process diagram that demonstrates the activities flow in Wikipedia training sessions (see Figure 5.18). This was presented to the participants in the second activity of the second session, "Define", who were asked to review this activity flow and identify any gaps or misrepresentations of activities. They were also asked to rate activities based on their views of how the trainees perceived an activity by labelling it with emojis (happy, neutral, and sad). This helped capture the trainers' feedback on their experiences with the current writing activities and establish the groundwork for discussing contingencies, as well as what could be done differently, had real-time group collaboration been introduced.

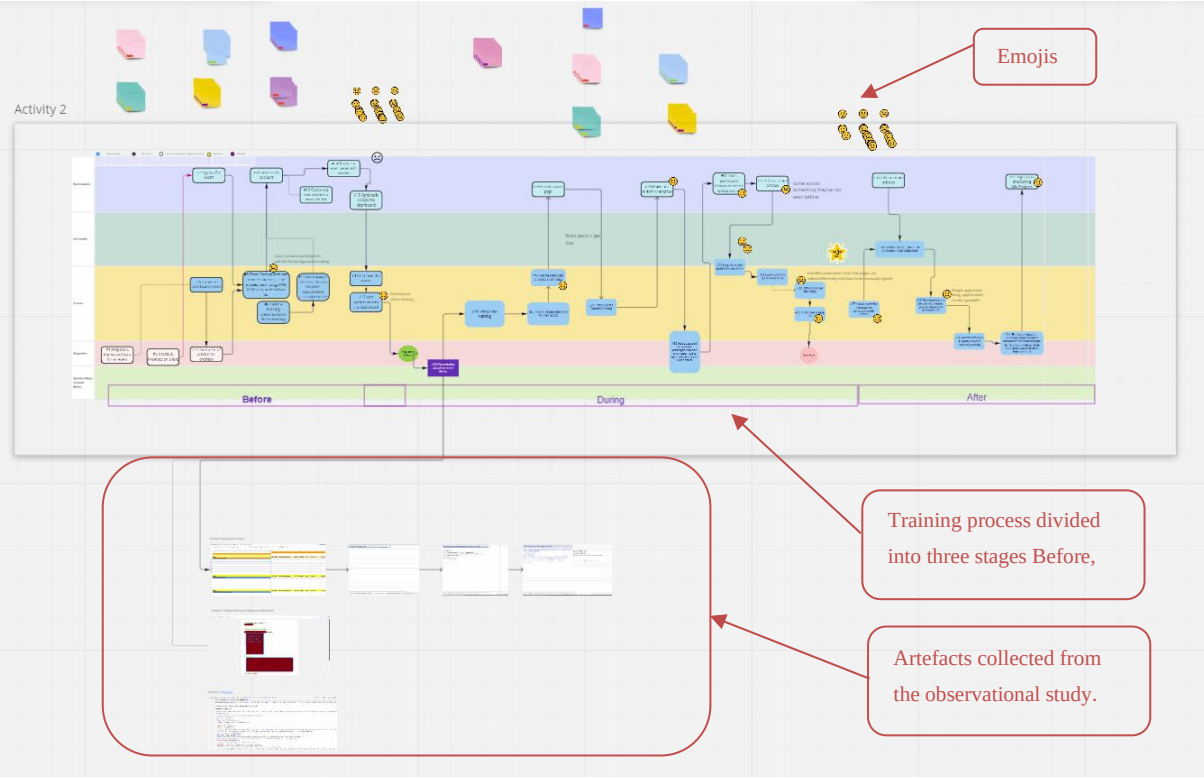


Figure 5.18 Training Activities Flow Review part of the Miro board in the "Define" session.

Define: Reverse Brainstorming

The third activity of the "Define" Session included reverse brainstorming on all factors that could prevent real-time group collaboration editing during Wikipedia training sessions (see Figure 5.19). Brainstorming as a term for generating ideas was coined by Alex Osborn in the mid-1950s (C. Wilson, 2013). Its fundamentals focus on the quantity rather than the quality of ideas, thus postponing judgement on the quality and encouraging thinking big and wild ideas.

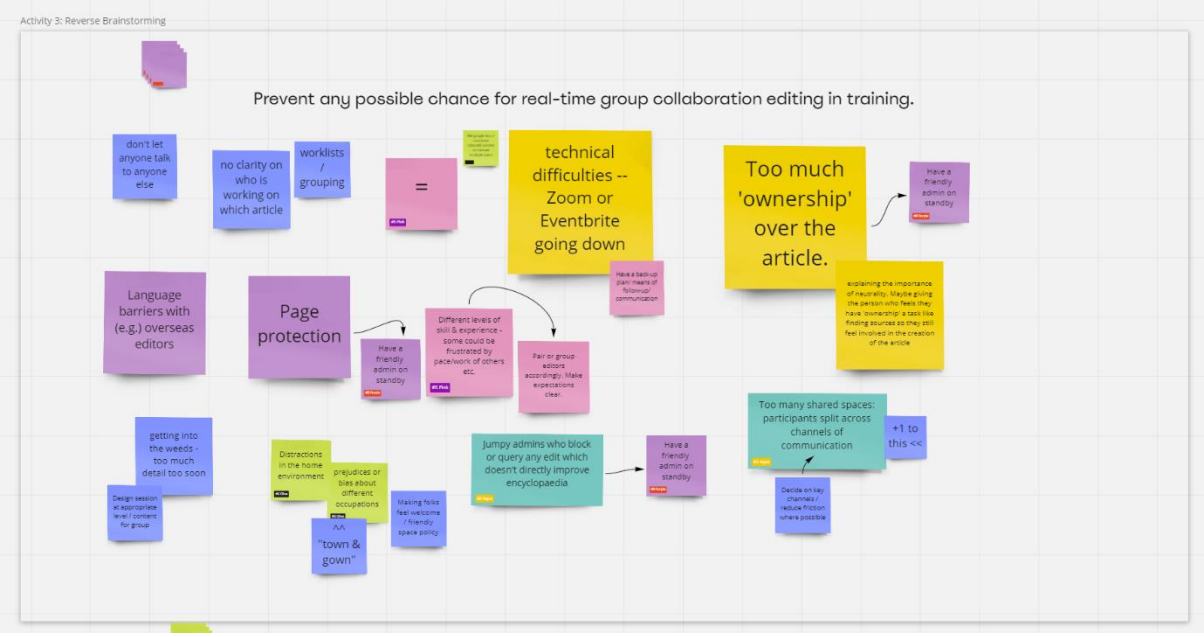


Figure 5.19 Reverse brainstorming activity part of the Miro board in the "Define" Session.

In a variation of this method called reverse or negative brainstorming, a facilitator asks about the negative facets of a topic. The logic behind this is that it is easier to find the faults of something and then use these inputs as a starting point for finding a solution (C. Wilson, 2013). For instance, instead of asking how to make a seamless user experience for an application, one should ask how to make the application unusable or as irritating as possible for users.

Discussing the introduction of real-time collaboration can be uncomfortable for many in the Wikipedia community due to its contradictory aspects with the community norms and traditional ways of collaborative editing. Therefore, as discussed later in 5.5.3, this research has found that using this tool of reverse brainstorming helps make the community comfortable criticising the feature and later find reasonable solutions to critiques.

As part of the reverse brainstorming activity, participants were asked to review their contributions to the SWOT analysis activity in the first session, "Identify", which was placed on top of the activity space on the Miro board, as shown in Figure 5.20. This was to help them make use of the training strengths, external and internal challenges, and opportunities to brainstorm ideas that could prevent or hamper real-time collaborative editing in the training session. The final step of the activity was to find solutions to what could stop the use of real-time collaboration.

Figure 5.20 Full view of the reverse brainstorming activity part of the Miro board with the supporting information in the "Define" session.

Define: Co-designing a Wikipedia Training Session Activity

For the last activity of the "Define" session, the participants were split into two groups. Each was asked to collaborate on designing a Wikipedia training session that involves real-time collaboration based on the activities flow provided in the second activity of the "Define" session (see Figure 5.18). Participants had the option of ignoring the activity flow and starting from scratch with designing their training. This design activity included four main steps:

1) Designing a training scenario should:
    a. Be flexible to adopt different preferences for external tools,
    b. Involve real-time collaboration,
    c. Have a clear purpose,
        i. Why is the training needed?
        ii. What is the issue or opportunity?

    d. Identify the people involved,
    e. Define the participant's roles and background stories (age, education, culture, gender, work position, history, etc.). It does not have to be large; it could be as small as a trainer and two or three participants.
2) Pitching the training flow to other members of the session with an example of how it could be run.
3) Collaborating with the other group on merging the activities identified as strong points from each group's training design into one training scenario that could represent a broader range of participants.
4) Defining the pain points of this training scenario.

The first step is aimed at encouraging participants' thinking about concepts, use scenarios, roles, and user groups, while the rest are focused on "converging" the participants' design thinking, helping them find a shared understanding of a training context that allows real-time collaboration. At many stages of the sessions, it was emphasised that the goal was not finding the solution to introduce real-time collaboration that could be generalised in many usage scenarios but, instead, to find a solution that could be the foundation for an exploration phase that transfers its benefit to another usage context.

5.3 Analysis

Participatory design is characterised by its collaborative sensemaking (Teli et al., 2020). Part of the data analysis for the Discovery phase took place during the sessions in collaboration with the participants. For example, mapping the Wikipedia training activities, SWOT Analysis, and Reverse Brainstorming activities possessed an analysis element, whether linking challenges with solutions or analysing possible opportunities.

In addition, after each session, the participants were sent a thank-you note with an invitation to send feedback and comments on anything they could not say or do at the time of the session. The emails also had a link to assess the session with a Miro board designed for a "Rose, Bud, Thorn" exercise used to help understand what was good about the session, what was not, and the opportunities to improve the sessions. Only one participant used the board to leave a sticker with a happy emoji outside of the activity area.

After reviewing each session's data multiple times, the session was summarised, covering the main discussion points from the session and the direction of the PD process. This was then sent to the research collaborators for review. Once ready, it was sent to the participants, requesting edits, suggestions, and comments (see Appendix F).

This helped ensure that the understanding of the session and the summary captured the essence of each session. It also provided an efficient way to help the participants who missed a session make sense of what they missed and catch up with the rest in the following session. Each session had a summary of the previous sessions, too. For some sessions, two versions were provided: A long and a short summary version, to make the process inclusive for those having limited time who just wanted an overview of the research progress.

This continued for the first three sessions. However, Chapters 6 and 7 discuss how, as the study progressed, given the large amount of data, it was difficult to perform this analysis cycle. This led to using more advanced analytical lenses that helped extract insights to address the research question.

5.4 "Identify" Session – Insights & Findings

This section covers the session's key takeaways that will influence the design activities of the subsequent sessions. The insight section from each chapter gradually answers the research questions of this PhD research on introducing synchronous collaboration to Wikipedia.

Held in December 2021, session 1, "Identify", lasted for two hours. All three activities to characterise the ecology of Wikipedia training tools, mapping Wikipedia training activities, and the SWOT Analysis were completed, resulting in rich discussions and lessons that would be helpful in the following "Define" session.

The following section highlights the key takeaways of the "Identify" session.

5.4.1 Motives for Editing Wikipedia & Collaboration Experience

Participants mentioned many reasons for editing Wikipedia articles. For Participant D, this included the development of new "skills in writing, researching that they did not get from doing two degrees." and interacting with the global audience: "when you write there (Wikipedia), you know it's available to the widest possible audience". Participant D mentioned that "Wikipedia, and unlike a lot of like all the other media and online communities, it's so focused on knowledge" a place "you get the experience of making something better than anyone of us could make." And the fact that your writing "in English will be translated into other languages" is "really gratifying". The Wikipedia community was another reason for Participant B to edit Wikipedia, who felt it was "very welcoming, and it's easy for you to find your kind of place in it."

Several approaches to collaboration were mentioned when participants were asked about the way they collaborated with other editors on Wikipedia. For example, Participant A described a situation where community members collaborate asynchronously on the article

review process that involves interesting dialogues to improve its content to reach the featured level. The participants mentioned, “It’s a small part of what I do, and it’s a very small part of the whole of Wikipedia, but there, I think it is genuinely collaborative.” After being asked to expand on this collaboration pattern, Participant A stated:

> “The wiki software is designed around that [asynchronous collaboration]. It’s unlike things like Miro and so on. It’s designed to enable people to improve a particular page and make a series of improvements. There are some problems with it, particularly if you get people editing at the same time. You can edit conflicts and so on, etc., but the Wiki software is designed for collaborative writing of stuff.”

For Participant A, writing Wikipedia is built on a technology that mostly encourages asynchronous collaboration with little to no close real-time coordination or real-time collaboration. However, the participant referred to a few outliers where projects within the community could facilitate high-paced “collaboration to particular targets, whether that’s subject areas or types of problematic content.”—Participant C further described other similar collaborations in WikiProject, such as helping someone who enquired about their area of expertise. Subsequently, Participant D mentioned the collaborative effort in supporting newcomers in such projects as the Teahouse.[75]

Participant D linked the discussion about collaboration to questioning the extent to which Wikipedia is open for anyone to share her knowledge collaboratively. They mentioned that:

> “It’s open to everyone. And then they (Participant D is referring to a new Wikipedia editor experience) make a good faith change, and then the kinds of things that (Participant A) is talking about happen: They get the templates unfriendly; template message slapped on the article they’ve edited. They get an unfriendly message on their talk page, and they get this (newcomer referring to editing Wikipedia) isn’t open. This is baffling […] these messages are in a kind of language I don’t understand. I don’t understand what I’ve done, and so the openness is kind of theoretical. It’s not a practical openness to everyone [...], and that’s how I see my role as a trainer (Participant D, referring to themselves).”

[75] https://en.wikipedia.org/wiki/Wikipedia:Teahouse/About

The openness challenge of the current collaboration model motivated Participant D to be a trainer "to make the openness of Wikipedia a practical thing".

After a follow-up question focusing on the participants' previous collaboration experiences, Participant B responded by saying:

> "My kind of experience of collaboration not really worked jointly with other people in articles, but I believe the sense that I'm collaborating when I'm just doing small edits, and then that's recorded. And then other people will build on that. So, it's a collaboration that I don't see, but I know it's happening."

Reviewing these discussions, specifically, the last point made by Participant B, shows that the feeling of belonging to Wikipedia's asynchronous collaboration is developed through time and is primarily found in projects in which only core Wikimedia community members are involved. Most of those joining the community as newcomers do not have the opportunity to learn that collaboration does not happen until they get some of their edits changed or even deleted.

Also, the fact that Participant A became a Wikipedia trainer to address some of the challenges newcomers face shows that the community view the Wikipedia training context as a medium for bringing positive change.

5.4.2 External Tools Used for Wikipedia Collaborative Editing

The observational study already revealed the various tools used in the Wikipedia training context (see Chapter 4). The "Identify" session resulted in an expansion and a more in-depth discussion of this list of tools.

The participants used several tools for event management, searching for resources, social media, conferencing, and real-time collaboration. A case in point, they indicated using Google Docs, Google Sheets, Twitter and Etherpad to help them edit Wikipedia.

Participant D expressed this evolution in tool usage by saying:

> "Surprisingly, was having a social media back channel, so as I was leading Edit-a-Thon, which had some people in a room and some people remote. But they were all members of a Facebook group, so they could all post, so this was easy for them to make a post about what issues they were having or what state their draft was in. And they could contact experienced Wikipedians and get help. That wasn't an easy thing for them to do when maybe using the different talk pages and forums on Wikipedia itself would have been difficult."

This participant later justified this by stating that the users' prior experience plays a role in tailoring the training to the audience's background and setting. For example, those published in academic journals might believe there is no editorial board on Wikipedia to review the articles and that Wikipedia works entirely differently.

> "People who aren't experts or scientists have an experience of contributing to forums or Facebook or online discussion, which, again, works in a completely different way; you never have to sign your messages on any other site. You never have to put in a thing that says who left this message and when. That's what the software does it. And so, again, it's completely absurd too. People have this reasonable expectation about how these things work, which in the context of Wikipedia, is just wrong."—Participant D

Facebook or other online forum tools were used to ask more experienced editors questions about editing Wikipedia pages in the same group or even share drafts they had developed to be reviewed.

Lastly, Participant C described using WhatsApp as part of a Wikipedia session to handle questions and provide a backchannel for discussions among trainees:

> "If somebody had a query, they were just raising it through their WhatsApp group, and that helped as a back channel, and they were sort of self-helping each other. Because this was a remote session done during COVID, that was really helpful. In an in-person session, they would have just called across the room or sat next to somebody who could have shown them. But the WhatsApp group helped them."

Reviewing the tools led to discussions about the hybrid training model, where some attend in person and others join online, giving trainees flexibility.

This activity helped facilitate a broad conversation about the participants' collaboration practices, views on recent changes, and the emergence of new real-time ways of coordinating training and collaborations. Participants mentioned how trained groups opted to coordinate their tasks using easier and more familiar real-time tools, as instant messaging apps, or social platforms such as Facebook and WhatsApp. This emphasises the importance of answering the research question on how to introduce real-time collaboration for Wikipedia. Another finding was that most experienced editors assumed collaboration as an asynchronous practice, unless they were explicitly asked about the real-time one. This shows how the two decades of relying

on MediaWiki as an asynchronous tool built norms around it that are challenging to change. This underscores the importance of making change within a context, such as the training in which the community is more tolerant of experimenting with new tools and methods, rather than the broader application of real-time collaborative writing across Wikipedia.

5.4.3 Second Version of the Training Activities Flow Challenges and Opportunities

As shown earlier in Figure 5.11, the second activity discussed the different efforts, tools, or elements in Wikipedia training. Participants collaborated on listing the activities in chronological order, i.e., before, during, and after the event.

This activity complemented and sometimes changed the picture that the observational study had resulted in during the training process (see Chapter 4). Participants focused on the pre-training experience, which led to a discussion on the importance of confirming the users prior to the session and gathering information about the participants' needs and technical settings, e.g., available devices and Internet connection, to tailor the training accordingly. For example, the PC and Mobile interfaces of Wikipedia are quite different and, therefore, influence training and demonstration sessions.

Similar views were observed in reviewing the Wikipedia community discussion over introducing new features supporting real-time collaborative writing (see 4.3.3). This leads to further questions on making the training experience more flexible to fit the needs of different groups.

Therefore, this session's activity resulted in an iteration of the initial training flow diagram that was created based on insights from the observation of Wikipedia training sessions, as well as the participants' insights from Session 1. Section 4.4.4 shows the initial flow diagram. The new one after the iteration is divided below into three separate diagrams for clarity. The major changes include: splitting the training flow into three stages: before, during, and after. Also, more activities were added. The session showed the influence of activities taking place before the training. This led to adding several activities to demonstrate what takes place before the start of the training and involves the coordination that occurred before the session. The activities placed on two lanes indicate the involvement of both actors. The co-trainer has been given a separate lane. Apart from these changes, however, the activity did not reveal any new actors other than those previously identified.

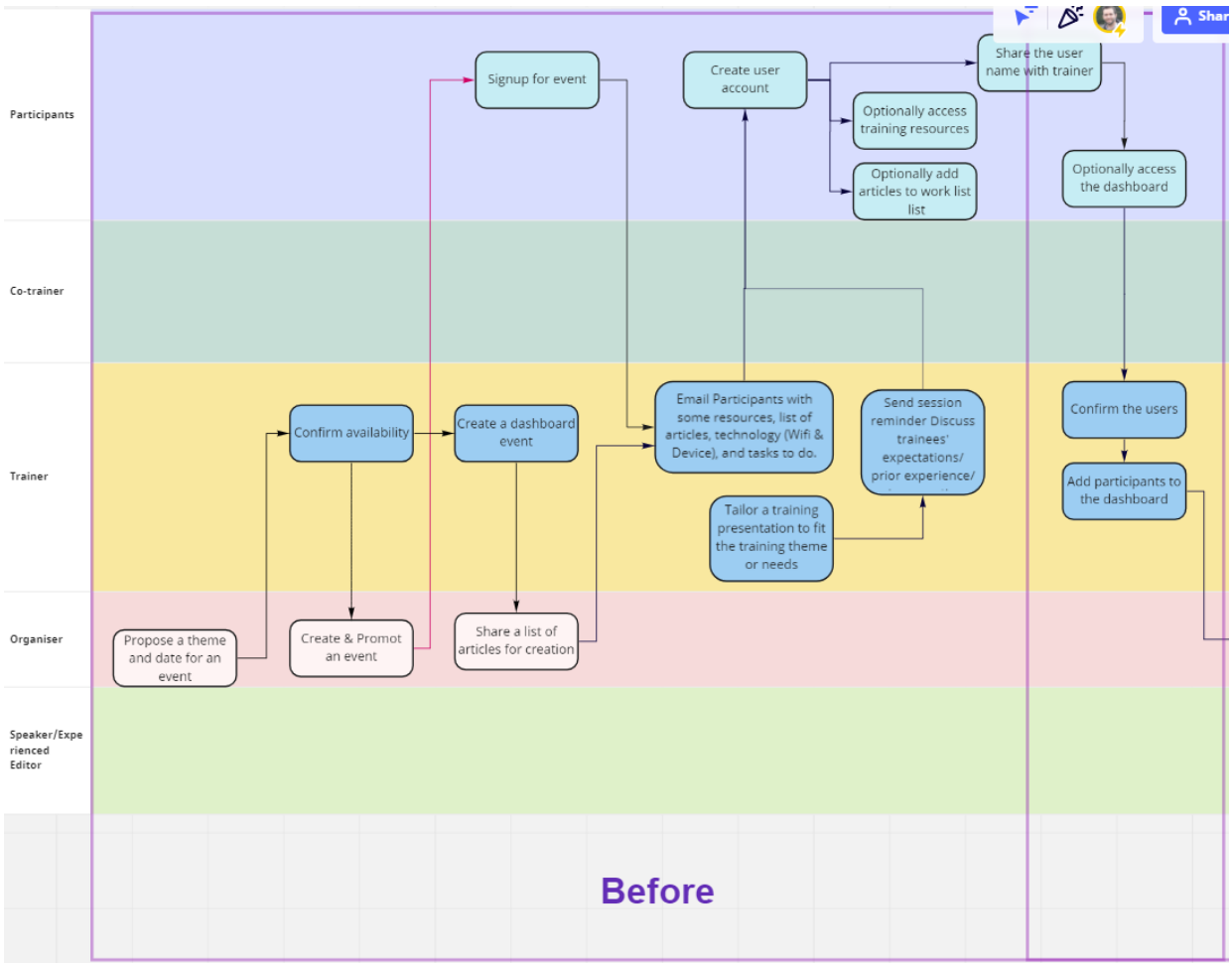


Figure 5.21 The "Before" stage at the new training flow diagram.

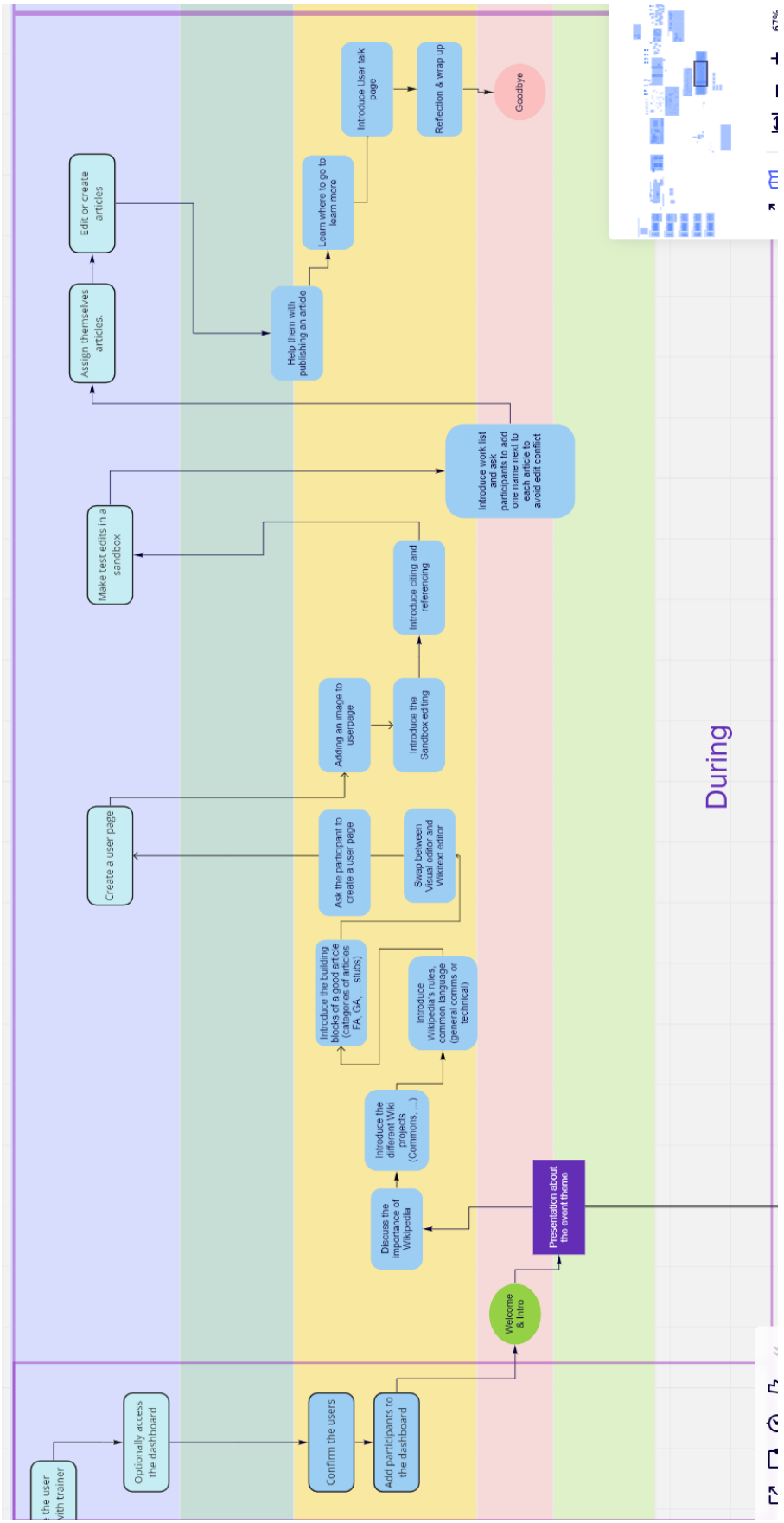


Figure 5.22 The "During" stage at the new training flow diagram.

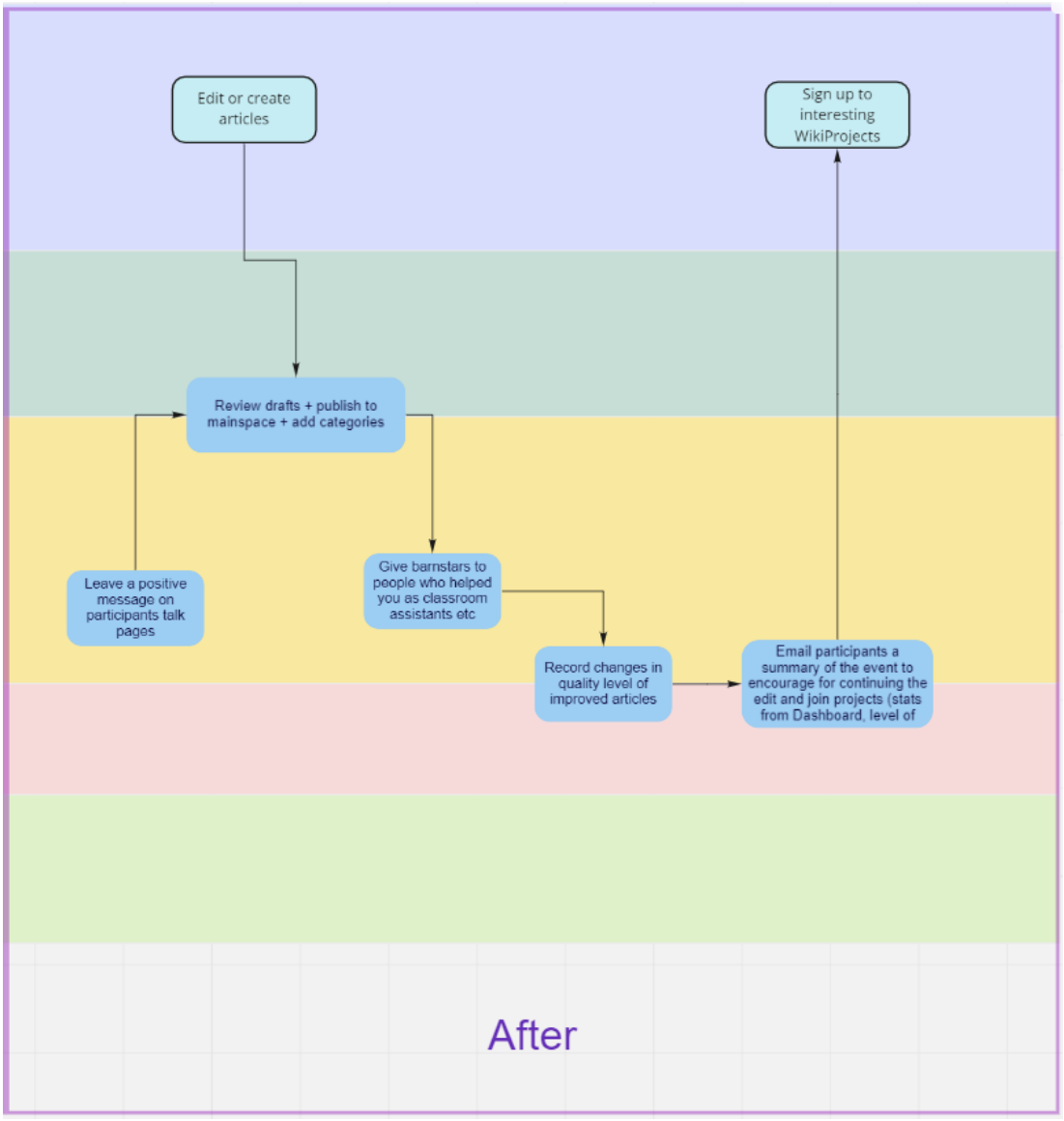


Figure 5.23 The "After" stage at the new training flow diagram.

5.4.4 Outcomes of the SWOT Analysis

As discussed earlier in Section 5.2.3, the SWOT analysis was split into two parts. The first discussed the strengths and challenges of Wikipedia training, while the second concentrated on the opportunities to improve it with a greater focus on collaborative editing and real-time solutions.

Strengths. Exploring the strengths showed a consensus on the importance of Wikipedia training as an efficient collaborative environment for trainers to share their hands-on experience and introduce people to Wiki editing and creating an article, helping them avoid the frustration of the openness that Participant D pointed out early on. In addition, they agreed that the training makes it possible for newcomers to directly impact their efforts in a shorter time frame than an independent onboarded process, generating higher gratification levels. On the

other hand, Participant E discussed how the feeling of collaboration and receiving thanks from strangers brought him joy:

> "I really like about Wikipedia, how some people can thank each other. I just think that is such a sweet kind of momentary connection with strangers who are grateful for something that you've edited, or some improvement that you've made. And it's something that's so simple to do as well. I just think that as a feature of the kind of interface, of Wikipedia is so nice, and it does feel like you're part of something and collaborating on something. Even if you are just writing a page and then leaving. If somebody is like thanks, you feel like you've helped."

General challenges. Participants identified certain challenges, varying from difficulties the new editors must deal with as newcomers, such as understanding the community terminology and conflicts of editing, to external problems, such as finding resources with balanced neutral views.

Technological challenges. Even though it was evident that most Wikipedia readers use smartphones, editing Wikipedia requires joining through a PC, which could be challenging.

Understanding Wikipedia editing conventions. It is also tricky to deal with the assumptions and expectations of some new editors about using the website. For example, any contribution to a talk page on Wikipedia should have a specific structure and signature of four tilde characters ("~") to show the commenter's identity. As mentioned by Participant D, this is unusual in all contemporary forums and social media platforms, which perform this function automatically.

Collaboration challenges. These included editing conflicts. A case in point is the conflict between trainees and established Wikipedia editors who edit an article simultaneously, albeit not in editing sessions with trainees. Participant E discussed an opportunity to solve this problem by developing an external tool specifically designed for work in groups on editing simultaneously before moving the text to Wikipedia. Such a tool should mimic tools that users may already be familiar with, such as Google Docs. Participant E attended a training where the trainees collaborated in real-time using Google Docs instead of the Wikipedia sandbox for drafting articles or writing a paragraph for an existing article.

> "So first they had the Google sheet which people could write against who [biographical article] they'd like to work on. But then they had people using the Google Doc to actually draft bits of articles and stuff, rather than using,

> you know their individual sandboxes necessarily as a starting point. So they used the Google doc more as a place to kind of gather information, and I guess also to build up something that could be used later on if people didn't get to create a full article. Or, you know, get as much of that done, which I thought was quite interesting. Although I think, I did inadvertently end up doing the same thing as somebody else, but we didn't notice. So maybe it's not certainly a flawless process. […] I think it is good. Something [Google Docs] that people are more familiar with, in order to help with writing or pulling content together and then building on that."—Participant E

This shows an awareness issue with real-time collaboration when participants cannot know who is working on what.

When asked about real-time collaboration, Participant A discussed a previous experience with a multilingual group that made sure editing conflicts would not occur by working on different language versions of the article instead of multiple people collaborating on the same one. However, Participant A did not indicate that the group collaborated on the article and only discussed how the group ensured editing conflicts would not occur. This was repeated by some participants who had their way of interpreting what "real-time collaboration" meant and avoided discussing it, as it would lead to edit conflicts, given the current asynchronous software structure they are accustomed to using.

Discussions in the "Identify" session later focused on new article creation to prevent editing conflicts with other Wikipedians, potentially editing an existing article simultaneously. If organised well, creating new articles would also give the trainees the experience of practising the creation of different parts of an article and the instant gratification of sensing the immediate impact of their training.

Participant B explained the "role division" in collaboration could play as a solution to the edit conflict, such as giving a task to someone to explore sources for an article and then sharing their information with the team.

> "If there's a small team collaborating together, maybe having assigned roles you know. So, somebody is maybe a timekeeper and somebody is kind of organising stuff. And then the people with the most knowledge on that subject can actually thrashing out the text, so to speak. You know. So, if people have different hats and different roles, so it might help them collaborate together."—Participant B

The benefits of having breakout rooms for collaborative writing sessions for smaller-scale collaboration were also discussed to reduce the distraction in the primary virtual training space.

Opportunities for Real-time Collaboration

This SWOT analysis in the "Identify" session helped reveal the possibilities and opportunities (see Figure 5.14) for training collaborative model improvements. Opportunities to adopt real-time collaboration were discussed from several perspectives:

Opportunities of Internal and External Tools. Some participants highlighted the opportunities of using internal Wikipedia tools to foster collaboration among participants, such as WikiProject, talk pages, and sandboxes. Others proposed using an external tool, such as Google Docs, for editing collaboratively before moving an article to Wikipedia. Another opportunity mentioned was creating a new tool that provides an environment with which the user is already familiar.

Opportunities as Part of the Collaborative Writing Process. As discussed earlier, many participants indicated the possibility of assigning different roles to trainees to facilitate collaboration. Dividing responsibilities into smaller collaborative tasks that could be performed collaboratively may lead to a better onboarding experience, especially for those who do not have the confidence to deal with all aspects of editing a Wikipedia article. For instance, Participant B suggested having a pair of one-to-one editing sessions, with one person responsible for researching resources for writing and another responsible for using suitable resources to write the article. Participant D reaffirmed the importance of exploring this opportunity by sharing with the participants a previous case when school children would be divided into smaller groups, with each group assigned a computer to follow up with the trainer.

> "Because not every student had a computer. They shared a computer each, and someone was nominated to be the writer. One person was nominated to be the researcher, and they'd come to the event with some online sources, so they divided their time up in that way."—Participant D

The SWOT analysis activity in my research has revealed many fragments and insights that point to several possibilities and opportunities for collaborative training improvements, as shown in Figure 5.14 for training improvements; moreover, several collaboration models were discussed. For example, Participant B suggested splitting people into pairs at the training writing stage, where one trainee would focus on researching content, and the other on typing and writing.

Participant A spotted a "big opportunity" in Wikipedia's Selected Anniversaries[76] **Leveraging Special Days and Anniversaries**[77] as a good opportunity for inviting people to collaborate on editing or creating articles for a social event. The participant linked this opportunity to a successful event where the trainees collaborated and coordinated editing tasks using a Facebook group a couple of weeks before an anniversary Edit-a-Thon.

Social Interactions as an Opportunity. The importance of increasing the social interaction between Wikipedia editors is highlighted throughout the "Identify" session. The importance of benefitting from talk pages to facilitate a discussion on coordinating work to improve certain articles was covered, and the participants mentioned leaving tailored messages on the talk pages associated with their Wikipedia trainees' user pages to celebrate their achievements.

5.5 "Define" Session – Insights & Findings

After the 2021 holiday season, a summary of the first session was shared with the current and past participants recruited for the "Define" session, which was run in February 2022. The focus was reviewing the revised version of the training flow diagram and discussing concepts and goals for a tool that could address some of the Wikipedia training challenges. The roles and tasks involved in running the training were also explored with the participants. The session covered designing training scenarios that included ideas from opportunities covered in the SWOT analysis in the "Identify" session, which was conducted for 2 hours using MS Teams and Miro.

5.5.1 Collaborative Editing Activities' Tools and Roles in Wikipedia Training

Based on an initial analysis of the contributions to the first activity, in the "Define" session, focused on the tools, three main categories of tools, as shown in Figure 5.16 , emerged. The tools are used for:

1) Planning training and coordinating the tasks.
2) Researching material for creating or editing the articles.
3) Tools used for collaborating on writing articles.

[76] https://en.wikipedia.org/wiki/Wikipedia:Selected_anniversaries
[77] https://en.wikipedia.org/wiki/Wikipedia:On_this_day

As shown in Figure 5.16, there was one interesting mention of a tool called "WikiProject Resource Exchange",[78] a Wikipedia page to help Wikipedians asynchronously collaborate on finding resources for articles.

Participant B pointed out emails as a tool used for coordinating tasks that were not included in the initial list of tools. However, Participant A explained that there were transparency concerns with using emails in collaborative editing articles, as follows:

> "Email is not transparent to people other than the two are communicating with each other. The only instance I can think of where it's common and tolerated as an activity on Wikipedia is if you're notifying an admin or an oversight so that something needs to be deleted, revision deleted, or oversighted. That's the only instance I can think of where email is currently recommended [...] we can use emails for organising events. But in terms of editing an article and making changes to something on Wikipedia. I mean, there's an assumption of transparency other than where you need to delete or raise something discreetly."—Participant A

Transparency is valued and required by the community in almost all collaboration activities. Mostly, this is intertwined with the demands for using open-source tools, which could be found in the community discussion forums favouring open-source tools with no corporation ties (*Vue.Js Adoption Discussion on Phabricator*, 2022). When the event was first advertised, and before starting the sessions, a Wikipedia community member emailed me, stating that:

> "The movement has a commitment to open source. That doesn't mean that we don't have collaborative projects running on Facebook et al., but it can be contentious."

On a different note, throughout the discussions in the "Identify" session, many attendees mentioned dividing the work of writing or editing Wikipedia pages and assigning different roles for a group of trainees as a worthwhile opportunity for exploration (see 5.4.4). This contributed to the design of the first activity in the "Define" session to explore this further.

The result of the first activity of the "Define" session (see Figure 5.17) offers a mix of roles and tasks. For example, having an experienced editor in a group of newcomers could help give real-time feedback on drafts as a "Reviewer". Some of the tasks mentioned were "Pasting pre-prepared documents into Wikipedia", "drafting/wikifying text", and "looking for other

[78] https://en.wikipedia.org/wiki/Wikipedia:WikiProject_Resource_Exchange/Resource_Request

places on-wiki to link to or from the page ". This thread was followed and led to future designs, as will be discussed in Chapter 7.

5.5.2 Activities' Flow Review and Evaluation

The second activity in the "Define" session focuses on reviewing the Wikipedia training activities' flow to identify any missing ones and evaluate potential pain points for trainers and trainees when conducting/participating in these tasks. This activity provides insights as to where real-time collaborative editing could play a positive role in addressing potential frustration or improving the positive aspects of asynchronous editing.

Participants discussed real-time collaboration activities, such as collaborating on a worklist ahead of the training session, to develop a sense of collaboration. However, the review and evaluation process led some participants to raise concerns about having high expectations regarding participants' engagement in pre-training events or readings. The participants also mentioned challenges in dealing with the Programs & Events Dashboard,[79] an external tool used by trainers to log the trainees' contributions after registering their usernames on the event page. During the event, the trainees are typically challenged by facing "something they have not seen before",[80] pointing to using the Wiki editor to write their first articles and engage with user talk pages.

On the other hand, participants mentioned that editing the Wikipedia sandbox increased the newcomers' confidence and made them happy about their progress. Activities that followed the training, such as rewarding editors who assisted trainers and encouraging trainees to register for interesting WikiProject with Barnstars (Wikipedia awards),[81] were recognised as a positive step.

5.5.3 Outcomes of the Reverse Brainstorming Activity

Reviewing the internal and external challenges in the first part of the SWOT Analysis at the "Define" session focused on the training session, on the whole. However, the reverse brainstorming activity was used to understand the challenges of introducing real-time collaborative editing in Wikipedia training sessions and propose solutions.

First, the participants were asked to review the challenges, strengths and opportunities that emerged in the previous "Identify" session and, based on that, expand the list of challenges that would prevent or hamper real-time group collaboration during article editing in Wikipedia

[79] https://outreachdashboard.wmflabs.org/

[80] This quote was added by one of the Define session participants as a text element on the Miro board flow diagram in Figure 5.18 next to activity #32 (Edit or create articles).

[81] https://en.wikipedia.org/wiki/Wikipedia:Barnstars

training sessions. Participants were then invited to revisit each challenge to collaborate on finding a solution using the Miro board.

Table 5.2 shows a subset of challenges linked to the activity they carried out and the proposed solutions from the SWOT Analysis Opportunities and Reverse Brainstorming Activities:

Table 5.2 Subset of challenges and possible solutions linked to the sessions' activities.

#	Activity	Challenge	Possible Solution/s
1	SWOT Analysis	Showing trainees the "Heffalump traps" and how to avoid them: Edit conflicts[82] Spelling conventions[83]	1) "Wiki comprehension" exercises ("under the bonnet") understanding the state of an article before editing. 2) Explain Talk page conventions - indents and four tildes ~~~~[84] 3) Nominate one person as the typist - everyone else verbalises input. 4) Google Docs-style editing tool.
2	SWOT Analysis	"Expectation that Wikipedia is very permissive and allows almost any kind of contribution, when the reverse is true." Getting over assumptions about Wikipedia (especially its reliability)	Out of this research scope.
3	SWOT Analysis	Users' prior experiences with publishing or discussion platforms work completely differently.	Google Docs-style editing tool.
4	Reverse Brainstorming	Confusing participants with too many shared spaces and channels of communication.	Decide on key channels / reduce friction where possible. Online breakout rooms for one-to-one coaching (less intrusive to the main session).

[82] https://en.wikipedia.org/wiki/Help:Edit_conflict
[83] https://en.wikipedia.org/wiki/Wikipedia:Manual_of_Style/Spelling
[84] https://en.wikipedia.org/wiki/Wikipedia:Signatures

5	Reverse Brainstorming	Technical difficulties with conferencing and registration tools.	Have a backup plan/ means of follow-up/ communication.
6	Reverse Brainstorming	Different levels of skill & experience - some could be frustrated by the pace/work of others etc.	Pair or group editors accordingly. Make expectations clear.
7	Reverse Brainstorming	Real-time collaboration in a training session might lead to unclear coordination regarding who is working on which article. This lack of awareness might lead to more than one group working on the same article or section of an article before sharing it on Wikipedia or with others.	Clearly define who is working on what on a real-time collaborative sheet like Google Sheets. Like Google Docs/Miro have coloured cursors to indicate multiple users.

Since openness is one of Wikipedia's core values, participants mentioned that the collaboration and discussions to form an article should be transparent for the public to view. Also, using a worklist that shows "who is working on what" is considered an important factor in successful collaboration. Having a friendly administrator who was aware of the collaborative efforts was regarded by participants as a possible way to address the challenges to be encountered by collaborative groups of newcomers.

5.5.4 A Scenario for a Synchronous Collaborative Wikipedia Training

As part of the last activity of the "Define" session discussed in Section 5.2.4, participants were split into two groups and asked to collaborate on designing a Wikipedia training based on the activities flow provided in the second activity. Due to time constraints, only one group finished most of the tasks involved.

They mentioned real-time collaboration activities, such as working on a worklist using Google Sheets ahead of the training to develop a sense of collaboration and screen-sharing for people to contribute ideas simultaneously, with one typing into an existing Wikipedia tool and the others discussing what to write. Something similar to solution number 3 for the conflict of edits challenge at the beginning of Table 5.2 in the earlier section.

Even though the observation showed that some trainers or organisers tended to send the working list before the training session, it was interesting to find that in one of the breakout

rooms, Participant E stated that she had never emailed participants working lists beforehand. The participant suggested the reason why working lists were not sent in advance using tools such as Google Sheets as follows:

> "I don't do it often because I've been worried about people picking up the same article or like accidentally deleting the Google Sheet, and I'll be worried with people deleting the Google Sheet or something when you're not there to supervise it. I think I wonder if that's just me being a bit controlling."

There are many interesting ideas to be discussed based on the scenario activity. For example, designing flexible components for facilitating real-time collaborative writing. In addition, this scenario might consider a larger number of challenges discussed in the Reverse Brainstorming activity and how the design fits the specific targeted groups' roles and background stories. The next chapter focuses on complementing this effort and positioning the solution with concrete, detailed usage scenario before outlining a concept for a system that could facilitate them.

5.6 Discussion of the Discovery Phase

5.6.1 Opportunities for Collaborative Tool Design

From the "Identify" session, it can be seen how Wikipedia training sessions depend on using different sorts of external tools that support real-time collaborative task coordination and editing.

It is challenging to identify the right assortment of tools for a specific group of trainees with diverse needs, as it hinges on several factors, such as:

- Trainers' familiarity and preference for certain tools and devices (i.e., conferencing tools and computer devices);
- affordability of tools and devices;
- legal obligations or subscriptions to specific tools from the organisation that hosts the training;
- the training context and themes. For example, a group of people who already know one another and want to be trained may not need advertising or an event management tool. On the other hand, training people to improve articles may require different tools, in order to start an article from scratch.

Reviewing the tools discussed and opportunities available seeks to find a tool that could act as the foundation for planning training tailored to the trainers' and trainees' needs and capitalise on new opportunities, such as real-time collaborative writing and task coordination.

As discussed in Chapter 2, real-time collaborative task coordination and editing have been featured many times on Wikipedia's community Wishlist and the Wikimedia Foundation planned activities. However, they stopped primarily due to the social challenges of introducing such a feature for everyone in Wikipedia. Nonetheless, as shown in the previous discussion, the online training set of tools already involves third-party real-time technology, such as Etherpad and Google Docs, as workarounds. However, the remaining challenge is introducing real-time collaboration by identifying a practical set of tools that must fit a wide range of groups and fulfil their underlying need for real-time technology as part of one system. Breaking down the challenge and introducing a solution for a smaller group of people in a relatively controlled environment is needed, which is covered in the rest of this thesis. By working with the Wikipedia community and newcomers on a solution that could facilitate real-time collaborative writing and task coordination, the first brick would be laid toward a new alternative collaboration model that might later scale and influence changes in scenarios other than training. This would have a long-term impact on consolidating Wikipedia's position as the frontier of peer-generated open knowledge.

Based on the opportunities discussed in this chapter, the next chapter's activities will include the following:

- Identifying scenarios that use real-time coordination and editing tasks. A scenario should describe training models with flexible components for facilitating real-time collaborative writing. It should provide a foundation for accommodating a flexible collection of external tools.
- Exploring the possibility of introducing a new tool that acts as a starting tool for trainers to plan training sessions that support real-time collaborative tasks coordination and editing.
- Identifying the user roles, and covering the who, how, and when real-time collaborative writing would be needed. User goals would be defined, which describe the tool's functionalities to benefit and motivate targeted groups of people to use it.
- Co-writing design guidelines to define a successful outcome. For example, defining what makes the tool relevant or attractive for the targeted group of people.
- Concluding with a short system concept statement.

D.4 in Appendix D covers reflections and lessons learned on conducting virtual sessions at the Discover phase.

5.7 Summary

Broad in its scope, this chapter is intended to capture interesting ideas. Having those in-depth conversations and activities with expert Wikipedians, some of whom possessed administrative roles, trainers, and Wikimedia Staff members, have led to:

- Defining the second version of training activities flow challenges and opportunities.
- Covering the challenges and potential of real-time collaborative writing in Wikipedia training.
- Identifying the importance of flexibility in training design.

This chapter already gives indicators of how the training scenarios should be designed. For instance, flexibility is important for real-time collaborative writing in Wikipedia training. In addition, co-designing a tool to facilitate such training requires methods to overcome the challenges discussed earlier in the reverse brainstorming activity. Finally, such a tool should be relevant or attractive for the targeted group from previous discussions, leading to a short description of a specific group perspective of using such a tool interface.

The next chapter, covering the Conceptualisation phase, reviews the outcome of the Discovery phase and continues with activities to identify detailed training scenarios that would benefit from real-time coordination and editing tasks.

Chapter 6 Conceptualisation Phase

Hagen et al.'s (2012) framework is divided into several phases, where phase 1, the ***Discovery*** and its "Identify" and "Define" sessions, explore community challenges. The ***Conceptualisation*** phase focuses on defining the user groups and collaboration process on a new system design concept, usage scenario, and guidelines. I describe the activities of the "Position," "Concept", and two "Review" sessions, in which participants were asked to collaborate on formulating opportunities to address challenges identified in the Discovery phase and co-designing personas and Wikipedia training scenarios that involve real-time coordination and editing tasks. This resulted in defining the concepts behind the WikiSync system that aims at facilitating real-time collaboration in Wikipedia Training (see 6.7.1).

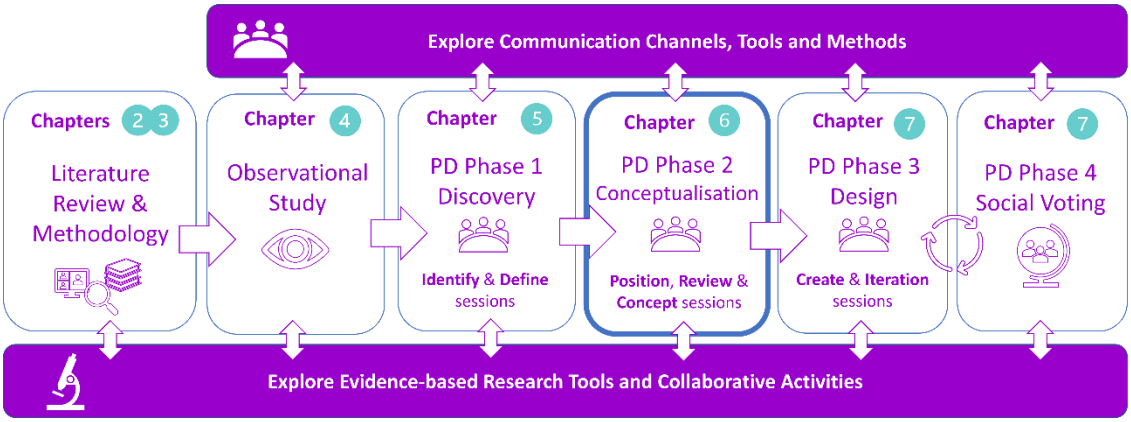


Figure 6.1 The Conceptualisation phase is highlighted in relation to other phases of this research study.

As shown in Figure 6.2, this phase includes a thematic analysis (TA) following the "Position" session that derived initial findings from earlier sessions to inform the "Concept" session, where participants reviewed and complemented the TA findings. All sessions relied on the same software and pre-session- and introductory activities used in Discovery (see 5.2.1).

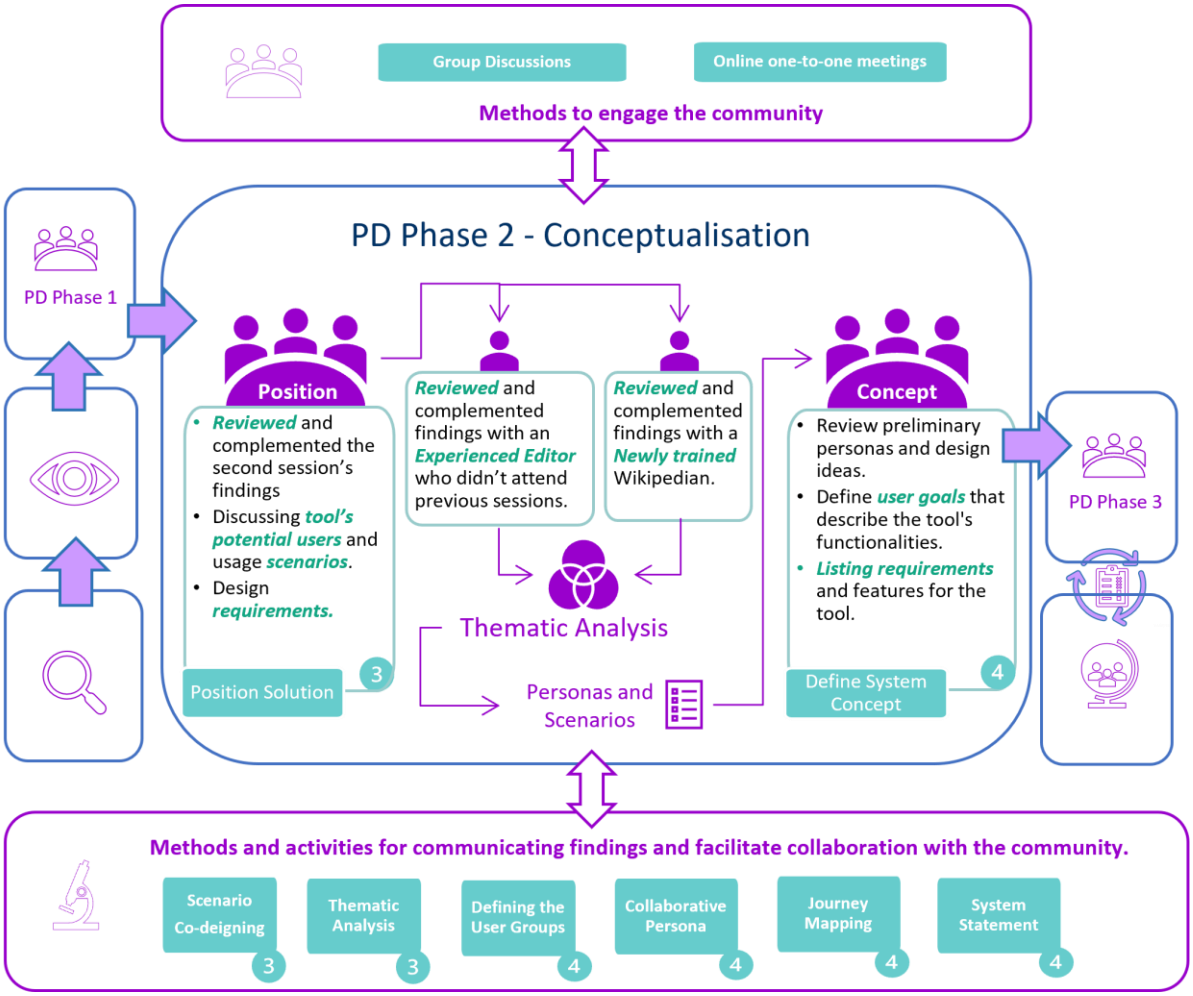


Figure 6.2 The Conceptualisation phase.

The Conceptualisation phase guides the system design process and answers **Q3 How can we co-design a tool and detailed training scenarios that would benefit from real-time collaborative editing?**

Answering this question results in defining a tool concept for a system built to enable Wikipedia trainers to tailor and deliver online training involving collaboration on writing Wikipedia articles in real-time.

Section 6.1 describes the data collection processes. Section 6.2 covers the sessions activities. Sections 6.3 and 6.4 describe the insights and findings leading to the thematic analysis covered in Section 6.5. Section 6.6 covers the "Concept" session findings. Section 6.7 concludes the chapter with a discussion of design considerations contributing to the Design phase.

6.1 Participants and Data Collection

This phase includes four online sessions between February and April, 2022: Two group discussions and two one-to-one sessions. The timing and dates were picked based on the poll data and the feedback from the sessions.

Of the 15 participants, 10 took part in this Conceptualisation phase, as shown in Table 5.1. The four sessions were audio and video recorded, and all statements made by participants were transcribed, resulting in 6.5 hours of footage, which were added to NVivo for analysis.

6.2 Conceptualisation Phase– Session Activities

The Conceptualisation phase consisted of 4 sessions, where the participants went through several activities based on the methods shown in Figure 6.2.

6.2.1 "Position" Session – Activities

The "Define" session generated ideas to explore real-time collaboration concerning training newcomers (see 5.5 and 5.6.1). The "Position" session would follow up on and complement this discussion. Participants were asked to outline the characteristics of groups that might benefit from a training scenario. Four participants took part, of whom three had participated in previous sessions.

Session Design & Procedure

The plan was to discuss training models that include real-time collaborative writing and cover methods to overcome the challenges discussed in earlier sessions. The focus was designing a workable scenario to involve the available tools, and workarounds mentioned previously to allow real-time collaborative writing in Wikipedia training.

The first part of the session involved splitting the attendees into two groups. Each group had two participants, of whom one took part in the "Define" session in the Discovery phase to ensure the continuation of the discussion. The "Position" session focused on positioning the new solution. Participants were asked to discuss the workarounds some trainers might opt to get their trainees to collaborate using third-party real-time tools. It covered questions that evoked participant memories from previous experiences to generate ideas for an alternative that substitutes the need for these workarounds. The goal was to envision the main features of a system built to support trainers in training newcomers to write articles in real-time. Before sending participants into their breakout rooms, they discussed the following points:

- The functionalities that benefit and motivate targeted groups of people to use the envisioned tool.

- How can the tool reflect Wikipedia culture and comply with its rules and guidelines?
- Design guidelines to define a successful outcome, including what makes the tool relevant or attractive for the targeted group.
- What type of articles can teams collaborate on?
- Who would benefit from such a system?

Figure 6.3 shows a section from the “Position” session Miro board (see Appendix D.5 for the full landscape) with a list of connected ideas produced by the second group to develop a training scenario. After finishing the discussions in the breakout rooms, participants returned to the main room to discuss their scenarios and define one that represents the best of each. The results are discussed in Section 6.3.

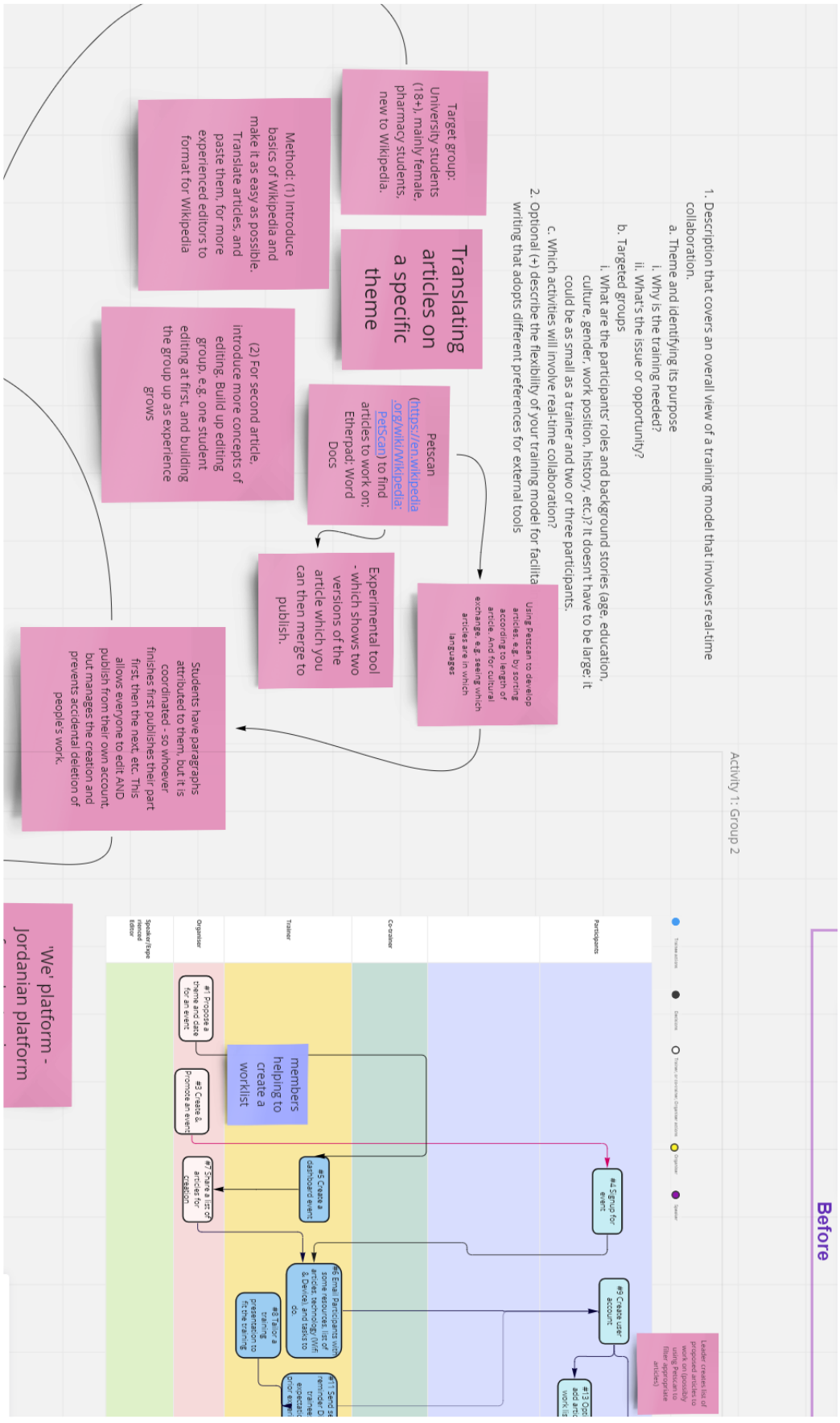


Figure 6.3 Part of the Miro board that shows the ideation that group 2 went through to come up with a training scenario.

6.2.2 Review Sessions – Activities

After the "Position" session, the need emerged to get diverse views of the findings gathered so far – from the Discovery phase as well as the "Position" session of the Conceptualisation phase. Therefore, as shown in Table 5.1, two individual review sessions were run with a trainer and a newly trained Wikipedian who did not attend previous sessions. The goal was to review the research direction, onboard new participants for the PD process, and complement previous findings.

Review Session 1. The first review session was conducted with an African Wikipedia trainer who did not attend previous sessions. I prepared a Miro board for her with a synthesis of findings from the previous sessions (see Appendix D.6). Moreover, the session involved open-ended questions, walking the participant through previous board activities, and a proposed real-time collaboration scenario that could be run using the WikiSync, tool being designed.

The discussion focused on complementing the research early results with the challenges Wikimedia faced in introducing a new collaboration model. The session discussed plans to get a broader involvement of the African Wikimedia community.

Review Session 2. The second review session was conducted with a newly trained Wikipedian. It provided a general overview of the research using the Miro board (similar to the one used in Review Session 1), addressing questions used in the "Identify" session (Discovery phase) with a focus on the newly trained editors' experience with collaboration on Wikipedia during and after the training session.

6.2.3 "Concept" Session – Activities

The data from the three sessions transcriptions were coded using NVivo to identify the main themes, the social and technical challenges, and the opportunities of introducing an alternative real-time collaboration technology. The finding of this analysis is discussed in Section 6.5.

This section focuses on the "Concept" session activities, where participants discussed potential usage scenarios and user groups of a potential collaborative editing tool. Eight participants took part, of whom four had already participated in previous sessions.

Session Design & Procedure

The session started with a presentation of the findings from previous sessions (covered later in Section 6.5). Participants were invited to ask questions and make comments. This was followed by an open discussion among participants about possible user groups that could benefit from

the first iteration of WikiSync. This discussion included brainstorming information about potential trainees and their personas that described their varying motives and levels of expertise as well as their needs as a group (see Figure 6.4). The concept of a collaborative persona is discussed in more detail below. A potential user journey was also demonstrated for the collaborative persona in this scenario. The session concluded with a review of a system concept statement laying out the proposed system's initial vision.

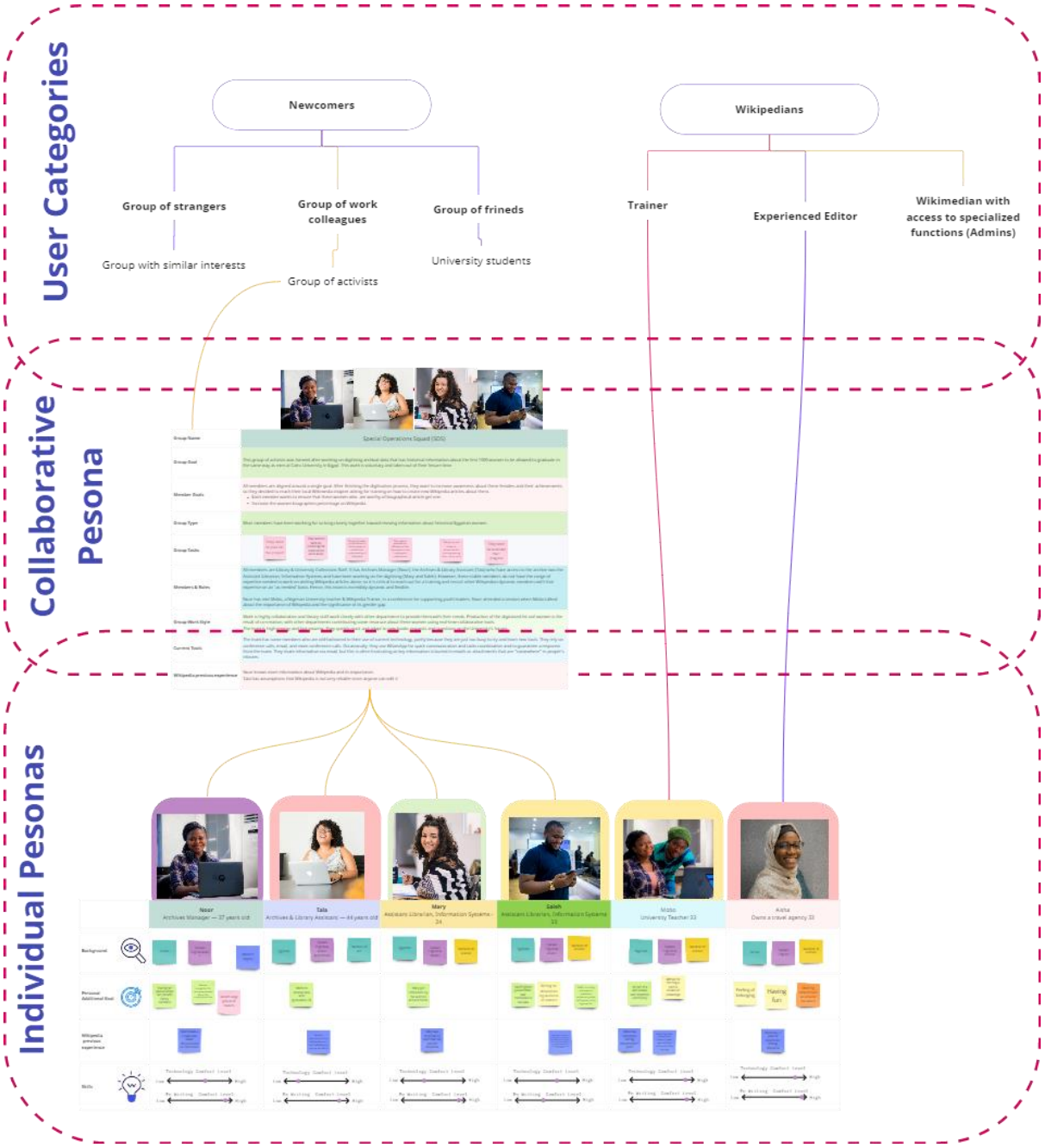


Figure 6.4 Defining the targeted user group activity at the fourth session.

Defining the User Groups. As discussed earlier, my research focuses on Wikipedia training as a suitable environment to bring change to the organisation. Therefore, the first activity of the "Concept" session presented several sets of potential users. These groups include characters that resemble what was described in previous discussions. The focus was on two main user groups (see Figure 6.5): experienced Wikipedians and newcomers. Each group was divided into subcategories. The first user group included trainers, experienced editors, and Wikipedians with access to specialised features in Wikipedia's MediaWiki software. The second user group included newcomers, which could include several groups of potential.

- A group of strangers with the same interest in contributing to open knowledge.
- A group of work colleagues who would like to be trained to contribute to open knowledge and advance their careers.
- A group of friends interested in improving a specific topic on Wikipedia and developing digital skills.

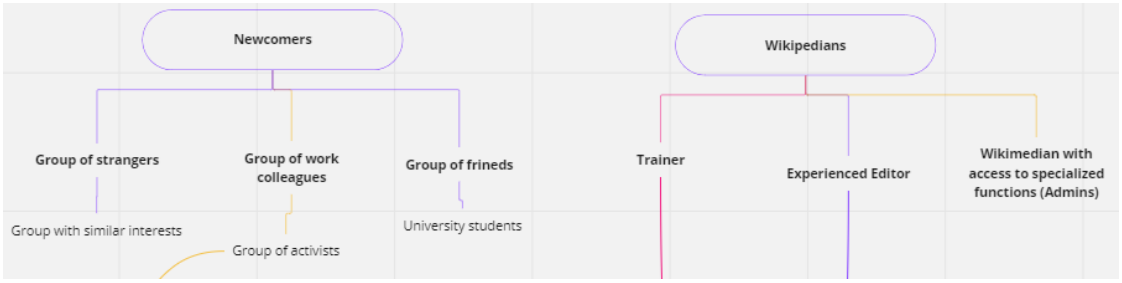


Figure 6.5 group of users.

The "Concept" session focused on a group of library colleagues who are knowledge activists. This group characteristics (or collaborative persona) were built based on the data that emerged from previous sessions before the start of the session. The group's story was shared with the session participants for review and collaboration on filling in missing pieces. The following sections discuss the reasoning behind this approach in more detail.

Collaborative Persona. The creation of personas is a useful procedure to help co-designers engage in sessions; according to Grudin & Pruitt (2002), such engagement highlights factors that would affect the user in adopting a new tool. Grudin & Pruitt show how a persona can help the designer have the end users' needs in focus, facilitating the creation of user-centred solutions. Also, developing personas can communicate the design team's assumptions about the targeted audience and facilitate decision-making. Furthermore, having a well-established usage scenario centred around personas can help the team memorise details that could otherwise be easily forgotten. Additionally, personas help communicate to the design

stakeholders the aim of the solution being developed, defining who the solution is proposed for and identifying how this would help stakeholders and users achieve their goals. However, Grudin and Pruitt (2002) raise concerns about using personas, as defining the right set of personas can be difficult, and there can be over-reliance on or over-use of this method instead of considering it as one of the user-centred methods used to design the solutions.

The new workplace demands for collaborative tools that capitalise on the advancements of Web 2.0 have increased (Matthews et al., 2010). However, adopting these collaborative tools is challenging. Matthews et al. (2010) have found that the group defaults to the traditional collaborations they are habituated to. They illustrate this by mentioning that some groups rely on collaborating on documents or file sharing by sending the document back and forth using email instead of collaborative tools. To solve this problem, the authors recommend developing group personas rather than individual personas to help them understand the group's needs and influence the designers to develop better designs that meet the various types of groups and communities. In a comparative study, Judge et al. (2012) find that collaborative (or group) personas to be a promising tool that "strikingly" achieves positive results, compared to individual personas, for example, leading to an extensive discussion and a focus on design features that support group interactions and promote collaboration.

Figure 6.6 shows a collaborative persona designed for a hypothetical group of open-knowledge advocates working together at Cairo University Library, Egypt.

Group Name	Special Operations Squad (SOS)
Group Goal	This group of activists was formed after working on digitising archival data that has historical information about the first 1000 women to be allowed to graduate in the same way as men at Cairo University in Egypt. This work is voluntary and taken out of their leisure time.
Member Goals	All members are aligned around a single goal. After finishing the digitisation process, they want to increase awareness about these females and their achievements, so they decided to reach their local Wikimedia chapter asking for training on how to create new Wikipedia articles about them. • Each member wants to ensure that these women who are worthy of biographical article get one. • Increase the women biographies percentage on Wikipedia
Group Type	Most members have been working for so long closely together toward reviving information about historical Egyptian women.
Group Tasks	They need to plan for the project They need to work on preparing the resources to write about The group needs to coordinate to find a trainer to solidify their understanding of Wikipedia They need to evaluate the efficiency of the time spent on the training and collaboration The group will have to advertise the training among their community They need to celebrate their progress
Members & Roles	All members are Library & University Collections Staff. It has Archives Manager (Noor), the Archives & Library Assistant (Tala) who have access to the archive two the Assistant Librarian, Information Systems and have been working on the digitising (Mary and Saleh). However, these stable members do not have the range of expertise needed to work on writing Wikipedia articles alone, so it is critical to reach out for a training and recruit other Wikipedian dynamic members with that expertise on an 'as needed' basis. Hence, this team is incredibly dynamic and flexible. Noor has met Mobo, a Nigerian University teacher & Wikipedia Trainer, in a conference for supporting youth leaders. Noor attended a session when Mobo talked about the importance of Wikipedia and the significance of its gender gap.
Group Work Style	Work is highly collaborative and library staff work closely with other department to provide them with their needs. Production of the digitaised list oof women is the result of co-creation, with other departments contributing some resoruce about these women using real-time collaborative tools. The team is high-energy and fast-moving. They quickly react and adapt to new books requests and questions on the Univeristy's forum.
Current Tools	The team has some members who are old-fashioned in their use of current technology, partly because they are just too busy to try and learn new tools. They rely on conference calls, email, and more conference calls. Occasionally, they use WhatsApp for quick communication and tasks coordination and to guarantee a response from the team. They share information via email, but this is often frustrating as key information is buried in emails or attachments that are "somewhere" in people's inboxes.
Wikipedia previous experience	Noor knows more information about Wikipedia and its importance. Tala has assumptions that Wikipedia is not very reliable since anyone can edit it

Figure 6.6 Collaborative persona for a group of activists who work together at a library.

More detailed descriptions of each character in the persona group who is participating in the fictional training, including the trainer and an experienced editor (see Figure 6.7), are discussed below, where the participants' demographics, backgrounds, previous Wikipedia experience and motives for learning Wikipedia are covered. In addition to the goals, group members' individual Wikipedia previous experience and digital and writing skills are also listed.

The collaborative persona describes a group of activists who formed a team called "Special Operations Squad" (SOS) after the archival data of historical information was digitised about the first 1000 women graduates from Cairo University, Egypt. Aptly, SOS work is voluntary and done during members' leisure time.

The team members share a **common goal**. Having digitised the data, they aimed to increase public awareness about those women's achievements. To that end, they connected with their local Wikimedia chapter in seeking training on creating new Wikipedia articles about the graduates, ensuring they were worthy of biographical articles, while raising the percentage

of Wikipedia women's biographies. Most members have long worked closely together to uncover information about luminary Egyptian women.

For the project of writing Wikipedia articles, the team must fulfil the following tasks:

- Preparing the resources to be covered in writing.
- Planning the project.
- Coordinating a concerted effort to find a trainer to solidify their understanding of Wikipedia.
- Evaluating the efficiency of the time spent on training and collaboration.
- Advertising the training in their communities.
- Writing the articles.
- Celebrating their progress and achievements.

The team's work style is highly collaborative in an environment that encourages working closely with other departments as they respond to their needs. For example, the digitalised list of women results from co-creation in coordination with other departments, including contributing resources about the women.

Brimming with energy and fast-moving, the team responded speedily, accommodating new book requests and questions on the university forum.

The team's **current usage of tools** varied from one member to another. Some were old-fashioned in their use of technology, partly because they were too busy to learn new tools, relying on conference calls and emails. Occasionally, WhatsApp is used for quick communication, task coordination, and to guarantee a response from the team. They shared information via email, which was frustrating, as crucial information could be buried in emails or attachments "somewhere" in inboxes.

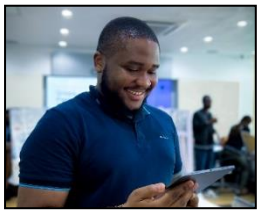
Participants' Demographics [85]

Trainee, Noor, is a 37-year-old Sudanese Archives Manager who speaks English & Arabic and has a master's degree in Archives and Record Management. Noor has high-level writing skills and is moderately comfortable with new technology.

[85] The participant's names are fictional, and the photos are under Unsplash license. https://unsplash.com/license

Trainee, Tala, is an Egyptian 44-year-old Archives & Library Assistant who speaks English, Arabic, and Armenian and has a Bachelor of Arts degree. Tala has elevated writing skills and is uncomfortable with the latest technology. Her motive for joining is to develop new skills (group work, IT).

Trainee, Mary, is an Egyptian 24-year-old Assistant Librarian-Information Systems. She speaks English & Arabic and has a Bachelor of Science degree. Mary has high-level writing skills and intermediate comfort with new technology. Her motive for joining is to celebrate women who have influenced her through their achievements.



Trainee, Saleh, is a 33-year-old Egyptian, Assistant Librarian-Information Systems. He speaks English & Arabic and has a Bachelor of Science degree. Saleh has below-average writing skills and is comfortable with the latest new technology. His personal goals and motives for joining the training are: (a) Saleh's great-grandmother is mentioned in the data; (b) writing for Wikipedia's large audience of readers; (c) Saleh runs a blog and wants to update his audience on what would happen to the digitised list.

Trainer, Mobo, is a 33-year-old Nigerian Information Systems teacher, librarian, and Wikipedia trainer, who speaks English & Yoruba and has a Bachelor of science. In addition, Mobo has high-level writing skills and is quite comfortable with new technology.

Expert Wikipedia community member, Aisha, is a 33-year-old British citizen who speaks only English and owns a travel agency. She has high-level writing skills and a high level of comfort with new technology.

SOS's Team Wikipedia Previous Experience:

Noor had a general idea about Wikipedia and its importance. Tala assumed Wikipedia is unreliable, since anyone could edit it. Saleh tried to create his grandmother's article, which was

deleted. Since then, he had not mulled editing Wikipedia, until Noor convinced him. Mary was surprised to learn that she could edit Wikipedia. Mobo has been editing Wikipedia for five years and only did online training. Last but not least, Aisha has 12 years of experience editing Wikipedia.

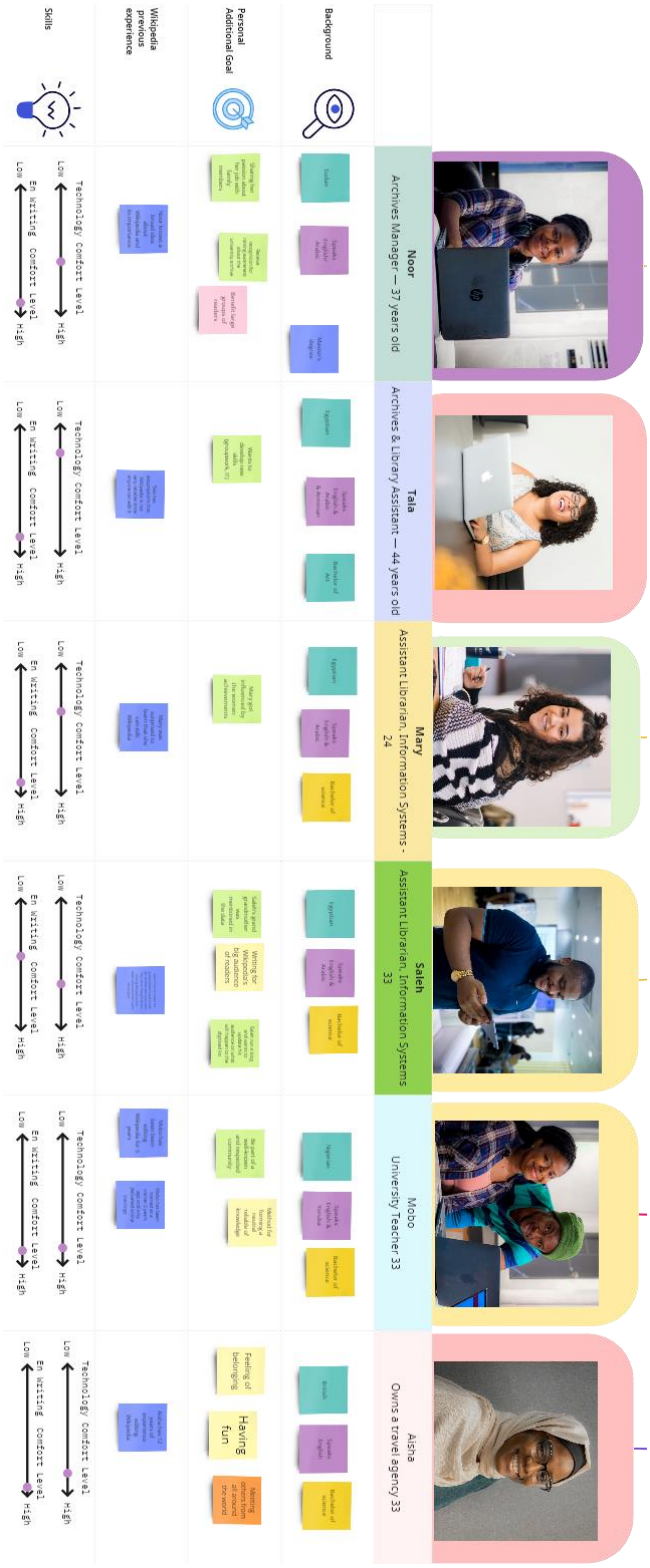


Figure 6.7 Personas of the training event participant.

Real-time Collaboration Training Journey Mapping. After going through the personas, participants were invited to define and discuss the user scenarios linked to the different personas (trainer, trainee, and experienced editor) that they could engage in to achieve their goals using a system that allows training with real-time collaboration. The focus was the key

interactions, goals, and possible pain points of the trainer, trainee and experienced Wikipedia editor as part of a real-time collaboration training journey. Figure 6.8 covers an early draft of the touchpoints of an onboarding phase of the trainer journey path with onboarding WikiSync. Additionally, the last two steps had some real elements of open-source technologies, such as Jitsi[86] and Etherpad, that could help build the tool.

[86] https://meet.jit.si/

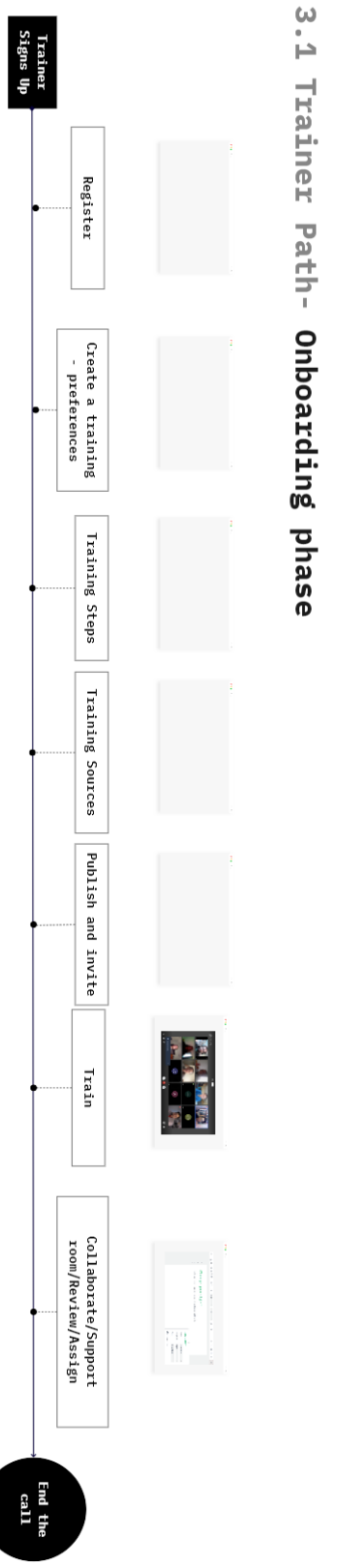


Figure 6.8 The onboarding phase of Wikipedia training using the proposed scenario.

I enquired the participants about their assessment of the user's sentiment toward this journey and how to improve it. Afterwards, we discussed how participants envisioned the look and feel of the system. Finally, they were asked to sketch their ideas using pen and paper, Miro, or unmute their Microphones to discuss with one another.

Defining a System Concept Statement. The previous sessions output and analysis helped shape the fourth session's activities. All stages of this session helped me draft a system concept statement, guided by the following questions:

- What is the system?
- Who is it for?
- What problems does it solve?
- What are its main features?

I provided the following system concept statement to my research collaborators for review and then to the participants in the last activity of the "Concept" session.

> "It is a system that will be built to fulfil Wikipedia trainers' needs to produce and deliver online training. It will offer an environment for groups of trainees to collaborate on learning, researching, and co-writing Wikipedia articles in real-time. The collaboration will be facilitated by an open-source conference tool (Jitsi) and text editor (Etherpad).
>
> Recent research shows that running an Edit-a-Thon could ease the challenges of newcomers and reduce the systemic bias in Wikipedia (Littlejohn et al., 2019; Gluza et al., 2021; Langrock & González-Bailón, 2022). Currently, Wikipedia trainee groups cannot edit a Wikipedia page simultaneously, and the learning curve of editing without training is hard.
>
> Thus, the new system will provide trainers with a flexible, transparent way to train small groups in real-time collaborative Wikipedia writing, starting with the training process and ending with sustainable volunteer engagement."

In a follow-up email, the participants were sent a video of the first part of our session recapping the study so far with the Miro board link, requesting them to access the session slides and comment on the system concept statement available through the shared Google Slides or a dedicated Etherpad page.

6.3 "Position" Session – Insights and Findings

The "Position" session activities helped co-design guidelines with promising prospects for successful Wikipedia training conducted by a hypothetical group using real-time collaboration.

The first activity involved reviewing the user groups, collaborative persona, and user journeys on the Miro board as well as splitting the participants into two groups to design training scenarios. Each group had two participants.

The first group discussed the complexity of having an external tool for text sharing to work on the articles since it might lead to edit conflict. However, drawing on earlier experiences, the group mentioned the benefit of a conference call that coordinates tasks and collaborates on different sections of an article to avoid conflict edits. The collaboration on a multilingual level was also highlighted to facilitate exchanging valuable resources among editors across different language versions of Wikipedia. The participants exchanged multiple ideas and scenarios about different possibilities for designing a Wikipedia training. They tried to complement each other's ideas.

Participant A, an expert Wikipedian with more than ten years of editing, was sceptical about the value of external text editing: "I'm not convinced that having an external, text-sharing thing is necessarily a good way to go in terms of collaboration." However, he described a real-time, collaborative experience he took part in, editing Wiki pages.

> "We had a thing come as a telephone conference on Skype where four or five we get together on that [...] That kind of work well certainly work for me [...] we had a few people on the screens across the northern hemisphere [...] We would pipe up and say, [...] 'Well, I'll go off and check this source' or 'I think, I added a category' or something."

This led Participant B to link the idea of group members doing different tasks to a previously discussed idea about distributing the roles of group members:

> "Sounds so similar to the roles thing we were talking about earlier with someone do research, and one person would do the actual typing."

Both participants agreed that collaborators should avoid collaborating on controversial topics as they attract endless debate.

This activity sparked interesting discussions and learning among the participants. The idea of asking them to develop a training scenario helped participant groups reflect on previous

experiences with a partial element of real-time collaboration and envision how such an external tool with real-time collaboration could help Wikipedians.

The second breakout room discussed real-time collaboration tools they had used before. Mutual learning was evident. Participant E praised Participant H for telling her about a tool that helps identify articles needing translation called PetScan.[87]

> "I haven't heard of PetScan before. So that's really interesting to hear about that and using it for translation."—Participant E

Their scenario focused on collaborating among students; similarly. They shared the idea of splitting an article into different sections for collaboration to reduce the chance of edit conflict. Again, Google Docs or similar editing tools were discussed to help edit and coordinate collaborative efforts.

More about the challenges

Previously discussed challenges reappeared in the "Position" session. For example, drawing attention to an article edited by too many people resulted in edit conflicts. The participants mentioned how the issue of two people submitting edits for the same article could be avoided with prior coordination. Another major issue was that real-time collaboration model in Wikipedia training risks excluding others from decision-making and the article's development process.

On a different note, Participant H mentioned the importance of gradually introducing trainees to Wikipedia editing, as learning how to edit Wikipedia from scratch is hard. It is also challenging to ask Wikipedia trainees to prepare for training beforehand. Leaving some of the work for the trainer was seen as a possible solution.

> "We just asked them to translate the article and just paste it. After that we fixed everything, we added source, we added pictures, and we fixed the page [...] But later on, when they had the second article, we started asking them about the resources, pictures and the other things".

At the session's end, Participant H described that the tool could facilitate collaboration by dividing the work into small tasks, such as improving the sentence structure and adding resources.

[87] https://meta.wikimedia.org/wiki/PetScan/en

A scenario for real-time collaboration

Considering the scenarios discussed in both groups, it was evident that they encouraged devising a way to coordinate tasks in real-time, whether by being physically present or having a conference tool for the groups to collaborate on an article with non-controversial topics. Both groups emphasised the importance of avoiding conflicts in editing which would discourage and complicate collaboration. This could be achieved by splitting the article among the group into several sections or working on different language versions.

Eventually, the possibility of a tool that facilitates organising collaborations over a conference call at a specific time was discussed. This collaboration should be available (transparently) for Wikipedia community members. For example, it could be posted as a message in the Wikipedia article's talk pages for collaboration or the Wikipedia community Village Pump. Additionally, the tool should allow users to book a section to work on to avoid a conflict of edits.

Again, the importance of finding a solution to the possibility of other Wikipedia community editors' work being stalled if we collaborated on live articles was discussed. One participant suggested it would be enough to let other collaborators know that, but later it was decided to narrow our focus to creating new articles, avoiding a conflict of edits complexity.

In the end, a tool's compatibility with a mobile phone was highlighted as an important requirement.

Edit History

Participant H indicated that he used to ask those who wanted to collaborate on the same article to work on a specific section that was then individually submitted to the sandbox. This helped preserve the history of published edits along with the article when moving it from the sandbox to the main space, leading to a partial answer to how to protect the history of edits from giving attribution.

6.4 Review Sessions: Insights and Findings

In the "Position" session, defining a concept tool to support real-time collaboration in Wikipedia training was almost reached. However, review sessions were needed to diversify the views contributing to the early stages of the design process and onboard new participants, whose participation and contribution were essential to solidify and complement our results and research direction.

6.4.1 Review 1: Insights from The Review Session with a Trainer

This session discussed findings from previous sessions ("Identify", "Define", and "Position"), trying to iterate on the ideas discussed or fill in any missing pieces. The session was supposed to include several participants who missed the early meetings. However, out of the five people invited, two accepted the invitation—one male administrator, a trainer from the Middle East, and a female trainer from Africa with three years of experience in Wikipedia. Eventually, only one participant from Nigeria would attend the session. The following are the main insights from the first review session:

African Wikipedian challenges

The session helped uncover a more diverse view of the increasing challenges faced by trainers trying to onboard new trainees. This confirmed that co-designing a tool with a better onboarding experience is needed.

> "I feel it's a good one [referring to co-designing WikiSync] and it's coming at the right time, because here in Nigeria we are faced with so many challenges."—Participant F

Previous sessions focused on the challenges the participants faced in editing due to the outdated interfaces, dealing with the Wiki tool and Wikipedia's rules. Further, the review session with Participant F revealed more challenges not mentioned in the early sessions. For instance, the participant mentioned how hard training new editors and retaining them was because the newcomers in Nigeria felt the Wikipedia community was "unaccommodating." The participant attributed this to the following:

- The many policies and guidelines for those joining to contribute with what they consider is limited knowledge about the topic. Also, she mentioned that "it's becoming very challenging [...] to get this new editor to understand this big community."
- Nigeria's official language is English; since some English Wikipedia administrators are not Africans, some of the African new editors' contributions to Wikipedia are deemed not notable enough to be on Wikipedia. The participant attributed this to the lack of the administrator's understanding of new information and its worthiness to be part of the English online encyclopaedia and dubbed it "Systemic bias."
- The participant explained that allowing these new editors to co-create articles is another challenge. "And so, I feel this research will go a long way to not only profile solutions. It

would also come on bring forth recommendations and even work towards the use and adoption of teams' real-time tools."

- Most Nigerians do not own computers. However, since Wikipedia's mobile app offers limited capabilities, the mobile view of Wikipedia's browser has a Desktop view option with almost the same options provided on the desktop. Therefore, they edit through the mobile browser using the Desktop view option.[88] This was mentioned by the participant from the Middle East.
- The participant considered training collaboration on new articles a "privilege" and guided her participants on only editing existing ones instead. This is because the difficulty of creating new articles increases when the topic lacks notability.
- The participant identified Blocking IP addresses[89] as a challenge for new editors who try to add or edit new articles only to find their IP blocked from editing Wikipedia. This happened with standard messages left on the new editors' talk pages with little explanation of the decision, leading them to ask the trainer for help as they "are scared and don't want to do anything."—Participant F

The multifaceted concept of transparency in Wikipedia

The second session's findings (see 5.5.1) discussed how Participant A emphasised transparency. Discussion with Participant F highlighted three dimensions related to transparency:

- The ***first*** one adds another layer of importance for Participant A, pointing out the necessity for trainees in our training scenario to be transparent about developing the article. According to the participant, this helps "the global Wikimedia Community's experienced administrators [...] ensure that members create reliable, verifiable information for others to consume."
- ***Secondly***, our proposed scenario from the "Position" session emphasised the importance of increasing transparency for collaborators or reviewers who could access the tool from the Wikipedia Community through a link provided before training. It provides clarity on the article's development by showing who is working on what. The proposed idea discussed in the "Position" session involves giving the trainees visibility of the Wikipedia

[88] https://en.wikipedia.org/wiki/Help:Mobile_access
[89] https://en.wikipedia.org/wiki/Wikipedia:Blocking_IP_addresses

community through links to community discussion spaces relevant to the training theme, allowing them to request help or collaborators.

Participant F pointed out the need to be careful about making the community transparent for the trainees. The participant reflected on the ideas discussed within the WikiSync tool scenario, saying:

> "I want us to look here at the user interface matters a lot. It's a big deal because, for example, most new editors can't even easily navigate through Wikipedia, and now we are talking about creating a real-time collaborative tool that will help, incorporating not just Wikipedia, but so many interfaces. So many menus like we got to get a space for a conference call and a space for people to send messages. We need to communicate with the global community, so now we also need to look at the user interface. Through that, we want to create. How easy is this tool for editors with no experience? They are just coming in. How will navigating through this tool be easy?"

- ***Lastly***, according to Participant F, the administrators should be as transparent with their decisions.

> "There's this feeling of new editors and even some experienced edits when they want to create a new article for the first time and within one hour or let's say you worked on an article for almost 24 hours, and at the end of the day someone just walking, just scrolling as an administrator, and you get a notification on your talk page. And within three days, or let's say, a day. The article gets pulled down by some of these administrators. So now I also want us to look at the aspects of this administrator. Why do they just get to pull down articles? Why can't they also create a solution by identifying the possible areas that need improvement and if possible, even providing reliable sources [...] how transparent are they in the mode of discharging their rules or their responsibility? How transparent are they themselves? We also need to look at that area."

Insight. This contribution helped define transparency differently, shaping the tool's design in terms of simplicity and group awareness.

Assumptions about real-time collaboration

After I walked Participant F through the WikiSync usage scenario of the "Position" session, her Internet connection dropped multiple times during the review session. This worried her, an example of how an Internet connection can be unstable.

Insight. This showed our scenario was built on the assumption that the defined user group would have a stable Internet connection. In the "Define" session (see Figure 5.19), one participant mentioned that the trainer might face technical challenges interrupting the process, such as a conference call service such as Zoom having lags or freezes during meetings. Earlier sessions did not consider trainees with poor connection, as that would hamper real-time editing. This might have resulted from the fact that in the previous session, most participants were located in the UK, where the connection is better.

This gave me insights into how to factor the element of inclusivity in the future WikiSync design, which offers fall-back features. For example, features could support those who dropped from the call to catch up with the rest or even contribute to tasks that would complement real-time collaboration asynchronously, such as gathering sources.

Opportunities for real-time collaboration

Since the focus is real-time collaboration rather than technical challenges, I did not want the Internet connectivity issues to drown out the discussion of real-time collaborative editing. Therefore, I explored with Participant F the viability and opportunities of a scenario which presumed people had an Internet connection allowing joining training and, subsequently a real-time collaborative writing session. To that, Participant F replied, "if there's no challenge having the Internet connectivity or bad Internet, I think that's a good initiative." She welcomed the idea and discussed many benefits she saw in the real-time feature added to the collaborative writing process after a training session:

- It helps people practice by collaborating on what they have learned.
- It helps reduce some of the challenges faced in creating articles and helps new editors "find their feet" when contributing to Wikipedia.
- It would make article creation "stress-free."
- It helps with the trainee's confidence "when they are making mistakes, somebody is somewhere guiding them on how to rectify some of those errors, So it makes them feel like, yes, this is teamwork."

- Finally, the participant added, “I see it’s going a long way to help provide solutions to some of the challenges I might not even be able to identify now.”

Proposing solutions and design recommendations

The same participant proposed the best solution from her perspective, which was mentioned as part of the “Position” session’s scenario as follows:

> “I think the best solution [...] is trying to get people to work together collaboratively, work on the drafts, maybe using Google Docs, or using some other collaborative tools. And getting somebody to [...] create that article on Wikipedia [...] I think that would work, and it will be more effective.”

The participant also proposed several ideas and amendments to the scenario for what the tool should be:

- **Giving people the freedom to choose their role rather than being assigned a role by the trainer**. The participant was worried that by not teaching people everything, including editing WikiText, we would be left with experienced editors who are unaware of the process.
    - **Insight.** Linking this to a previous discussion of how some newcomers with limited plans for contribution could be overwhelmed by all of the Wiki complexity led me to propose two categories of participants. One group would learn only how to contribute with their knowledge, while the other would contribute more frequently and be part of the broader Wikipedia community.
    - **Insight.** Furthermore, there is a need for a greater guiding role for Wikipedia administration through the proposed system.
- The participants suggested that “the tool should have samples of already created articles or a template guide on creating them [...] there may be a short video that can also be incorporated into this.”
- Having a target after the training would motivate the participants to use the tool.
    - **Insight.** This led me to say, “I could have conceived of another feature to help set a specific training target, as writing three articles by training’s end.” This notion was welcomed by the participant. More about this feature will be discussed later in Section 7.1.7.

6.4.2 Review 2: Insights from The Newly Trained Wikipedian Session

The observational study helped create a picture of the trainees' experience. However, there was a need to incorporate their perspective into the emerging discussions of the sessions. Based on previous sessions, this review session sought to find and improve on missing pieces from new editors' perspectives. Only one out of two who registered joined this review session. The female participant, an Egyptian living in Scotland, made simple edits after an online training seven months prior.

This trainee experienced real-time technology through collaborating on Google Docs and Sheets. She explained she was only a user who had been told many times that Wikipedia was untrustworthy, however, following training, she changed her view about Wikipedia.

The trainee had been told to make many edits before being allowed to start a new article. Then the participant explained her experience with a new feature proposed by the Growth Team to recommend newcomers' edits to get them started with editing gradually.

> "After the training, not exactly the same day, but maybe a few days later, Wikipedia itself recommended an article for me and this article, about a city I've never heard of before, even I didn't know that it existed. I tried to correct any grammatical errors. Maybe I did finally find something very simple to edit."

How would a newly trained person explain editing on Wikipedia to a group of friends who want to join?

After asking the participant how she would describe the process of editing Wikipedia to a group of friends, she said with little confidence, "I don't think I'll be that qualified to give them exactly how or explain the process." However, asked whether she would encourage them to collaborate on creating the articles, she recalled a time when she used real-time collaboration over a Google Doc, describing it as follows:

> "It's, of course, better because everyone had his point of view. I was able to feel this, especially when I was writing the [PhD] personal statement because I gave it to my husband to read and gave it to my mentor. Each one has his way of adjusting something or adding something in a different way. One has done a scientific review, and another has done the grammatical review, so

> together, they made things look better, so each has his own experience with something."

Asked about how to organise the collaboration for the Wikipedia trainees, the participant said:

> "It mainly depends on the people who are going to do this maybe and their experience as well. I would recommend, at the beginning [...], giving each one a part of the article to write or review first and then after they finish each one to have the whole article to read and review maybe a more generalised for everyone can have his own view, or on the whole article."

This statement was similar to other suggestions from the scenario. Another idea the participant spoke about was similar to that of Participant F from the first review session. She said:

> "Sometimes to be open and very transparent, sometimes it may bring confusion to others [...] to see everything all thing happened write or collaborate to edit this. As you said, conferences for the collaborations and all of these things, maybe I found this a little bit confusing. To understand everything happening, maybe I'll need besides this something like a manual telling me exactly what happened or summarising, so I don't have to go through all of these things or get lost between everything."

Finally, the participant indicated that if joined by others in collaborative writing on topics relevant to their domain, she would feel more motivated and confident to contribute.

6.5 Thematic Analysis

After the review sessions, moving the research along with the scattered information on the boards had become more difficult. There was a clear need to identify and organise key features into a more structured view. Additionally, there was a need to report to the participants a clearer vision of what these sessions were leading up to and to push forward the PD process to the next session, where the system concept would be defined.

This section details the process of the analysis phase that took place after the review sessions. The analysis involved the data from the previous sessions, as shown in Figure 6.9, where the funnel is a thematic analysis guided by the data and my knowledge accumulated from the observational study and previous experience with the community.

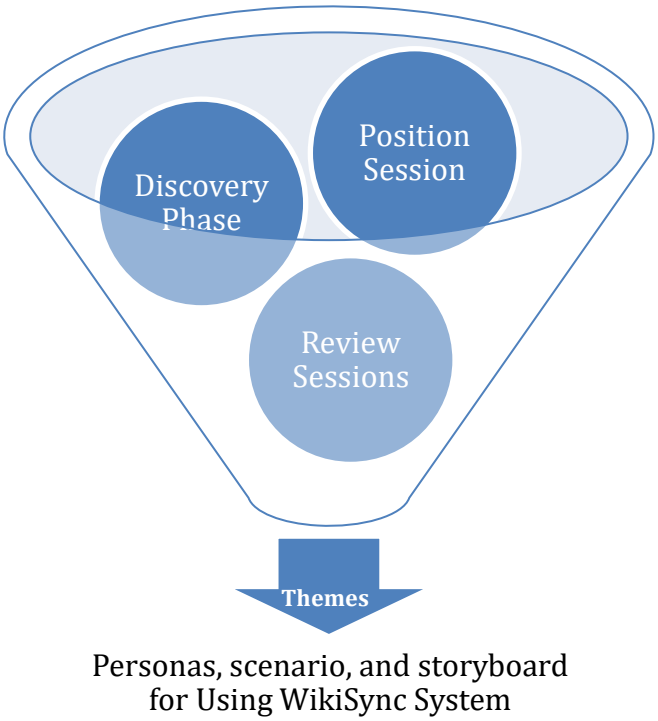


Figure 6.9 Using thematic analysis to produce themes that guide the design of personas and scenario.

This section illustrates the stages of data analysis to produce the findings, which were reported back to participants for verification and used to shape the "Create" session activities. The research design emerges in response to the needs during the research journey. In qualitative research projects, the "gathering and analysis of data occurs concurrently; it is a constant interplay of data and analysis, data informing analysis, and analysis informing new data collection" (Pickard, 2013a, p. 267). This interplay was apparent in the decisions shaped after finishing each session. I performed a small-scale analysis and regular reviews between the sessions with my research collaborators (see Figure 3.4). Each one contributed from their perspective and pointed me to specific tools or methods, helping me guide the direction of sessions and provide a concise summary for the participants. The small-scale analysis involved uploading the session video recordings to NVivo and doing an initial coding round to help identify the main findings or unanswered questions to pursue. This was afterwards sent to all participants to verify and comment on, which helped me design the next session. However, after finishing the review session, the need emerged to compile the data generated so far into something that can help drive the study direction to be presented visually to participants to avoid lengthy textual summaries. This led to the search for a qualitative analytical approach, helping me extract clear themes to represent the available data to participants and use it as the foundation for findings supporting the system's first design iteration. I had to take a step back from the process of facilitating individual sessions and instead look at the process as a whole

to give structure to what had been achieved up to that point as well as to focus subsequent sessions on themes agreed with participants.

There exist several approaches to analysing qualitative data. Well-known general strategies of qualitative data analysis are discourse analysis, thematic analysis and grounded theory, used depending on the context of the data and the expected outcome (Bryman, 2001).

Over the years, thematic analysis has become one of the common methods in HCI and CSCW to help generate concepts needed to design effective computing systems (Carvalho & Fabiano, 2022). Thematic analysis is widely used for "identifying, analysing, and reporting patterns within data" (Braun & Clarke, 2006, p. 79). Braun and Clarke define six phases that the analysis undergoes, which I utilised to structure the data analysis as follows:

Familiarisation with the data

In order to analyse the data, observational study notes were re-examined, and the five sessions were reviewed. I checked my notes about the sessions, rewatched the videos, and enhanced the automated transcripts and initial annotations.

Generating initial codes

The data from the Discovery phase, Position and Review sessions were entered into NVivo and coded along with the transcripts. I followed an inductive approach to coding.

After an iterative process of reviews and categorisation with my research collaborators, the codes list emerged gradually to constitute what is shown in Appendix D.8. Below, Table 6.1 shows how data segments were coded.

Table 6.1 Examples of data extracts and their codes.

Data extract	**Coded for**
"The biggest challenge so far for me is new editor getting messages from administrators on their talk page about their IP address being blocked, editing conflicts and so on like that"—Participant F	Onboarding experience -> Wikipedia Training-> Trainer challenges

“So it could be really good having multiple people working on the same thing with different levels of skills, obviously because they can help each other, but depending on the people, it could be, I guess, challenging. If you’re really new and you feel like things are moving a bit too fast for you or people are doing things in a way that you don’t understand or vice versa [...] if you just want to get on and do something and you can’t.”—Participant E	Real-time collaboration -> Challenges -> Different levels of skills.
“Play to people’s strengths - some like to write new stuff, some prefer to edit others’ stuff, others like to research sources, etc.” - Sticker contribution to the first session’s Miro board.	Roles and distributing tasks among the group.

Searching for themes

After three rounds of coding, I reached the point where specific codes could be brought together into identifiable themes that connect ideas and topics that recurred throughout the PD process relevant to the research focus.

Reviewing the themes

The set of candidate themes underwent many reviews by the research collaborators until a list of 9 themes emerged to describe 249 codes. The themes shown in Figure 6.10 depict those that emerged from discussing the co-design of a training tool to support real-time collaborative writing.

Figure 6.10 Final thematic map, showing nine main themes.

Defining and naming themes for participants

This stage involved refining the themes and building the overall story around them as a means to facilitate communicating it to participants. Braun & Clarke (2006), emphasise the importance of defining the "essence" of each theme and the roles each plays in the overall story, which was reported as a PowerPoint presentation.

Producing the report

There was a need to link the story to the reason why most participants joined the research: advancing Wikipedia. After a summary of the main research question and the methodology used through presentation slides, I introduced my report on the thematic analysis using three main components:

6.5.1 What Drives Wikipedia Forward?

In light of the emerging themes from the data, my research aims to find out what drives Wikipedia forward to gradually address the reason for introducing real-time collaboration. The themes that emerged from the data produced a graph on how Wikipedia sustains its existence and growth. The graph shows five interdependent building elements of Wikipedia. The components, including Wikipedia, are depicted as gears.

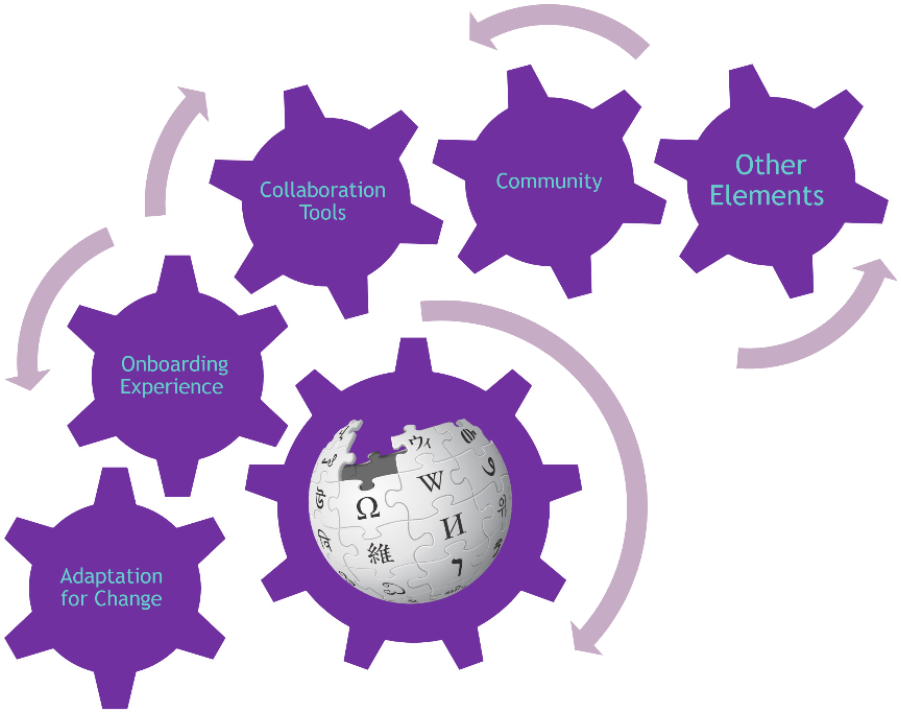


Figure 6.11 The emerging themes of what drives Wikipedia forward.

- "**Community**" gear represents the community backing up and contributing to Wikipedia.

- "**Collaboration Tools**" gear symbolizes the tools that enabled such a large-scale, synergetic, remarkable community collaboration.
- "**Onboarding Experience**" gear shows the condition for this community to proceed and thrive. We need a sustainable influx of newcomers who enjoy their first user experience and turn into fully-fledged Wikipedians.
- "**Adaptation for change**" signifies the importance of adapting to the ever-changing world to sustain Wikipedia's global leadership position as an open knowledge producer and provider, advancing its processes and technology for both readers and contributors.
- "**Other elements**" gear indicates other aspects not covered by this research that contributed to Wikipedia's global presence. However, the core elements are the community behind the collaboration tools used to achieve such results.

Expanding on the elements

The presentation included slides that expanded on sub-themes for the previously discussed elements. The slides are included as a story concerning the challenges and opportunities that Wikipedia training provides. This led to the need for change to meet the readers' and Wikipedia community's needs, followed by the opportunities we had, and finally, focusing on the asynchronous and real-time collaborative tools, exploring the social and technical challenges for introducing real-time collaboration.

The onboarding experience challenges.

Many discussion points showed how challenging it could be for someone with no prior understanding of who and what is behind Wikipedia articles to understand how it works. Here is a summary of the main challenges:

- Identifying the right path to becoming a fully-fledged Wikipedian.
- There is a lack of resources as some people interested in joining lack the time to learn everything needed to contribute. However, some lack the expertise, skill, Internet connection, motivation, or reliable content that they could use in building or contributing to Wikipedia.
- Bias leads to content being reverted[90] after spending time learning to edit and research.
- Poor user experience among the growing number of mobile users who edit through the desktop view or with the mobile app's limited capabilities.

[90] https://en.wikipedia.org/wiki/Help:Reverting

Wikipedia Training

On the other hand, Wikipedia Training's positive role was clear from the data. Among many other benefits, the training was seen as:

- A space for exchanging knowledge and expertise.
- A way to change assumptions about Wikipedia's reliability.
- An easier learning path that introduces the newcomer gracefully to the wider community behind Wikipedia.
- A means for newcomers to achieve a direct impact that brings instant gratification after publishing their edits at training's end.

Changing Needs

Discussions tackled the changing scenery of collaboration and the ecosystem surrounding Wikipedia. Newcomers became used to collaboration tools differently than those in the community who had used them for two decades.

Change in the way Wikipedia operates does not guarantee good results, as Wikipedia faces new, constantly changing, complex challenges as it grows more prominent, and the demands for its role as a neutral source of information are rising. Therefore, exploring opportunities that could bring constant positive change is essential, especially with its innovation process.

We noted how Wikipedia online training had brought new opportunities. For example, previously, some in remote areas could not access face-to-face training as travelling costs were high. Therefore, having the opportunity to reach out to them through online training was a great experience.

Another point of discussion was the opportunities for tailoring the training. This could help design sessions to fit different skills and tool preferences since not everyone is used to the tools we might provide.

Collaboration tools

Here are some insights that stemmed from the analysis of Wikipedia's default asynchronous collaboration:

- Throughout earlier sessions, when I brought up collaborative tools, participants assumed it was about asynchronous collaboration. Therefore, the asynchronous way was the defining concept for collaboration.

- There was a positive feeling about collaborating. The sense that one is part of a community that would develop what one contributes (see 5.4.4). However, it is not necessarily the immediate feeling one gets from the time one puts in your asynchronous contribution (see 5.4.4 and 6.4.1).
- Many discussed Wiki editing as a challenging experience, especially when it results in editing conflicts which have led many experienced editors or newcomers to lose edits. There was a discouraging effect of unmatching newcomers' previous editing experiences with contributing to Wikipedia options.
- Previous coordination with real-time tools such as Google Sheets yielded better results.

These insights linked to the challenges and coordination gradually led to the idea of real-time collaboration. The following are insights from discussing real-time editing:

- The participants were reluctant to discuss it. One of the main reasons was that everyone knew and was used to the idea that Wikipedia is like all other standard wikis that do not support real-time collaboration. However, many participants discussed how the community of trainers already used or accommodated demands from trainees' groups to use external tools, such as Facebook, Google Drive, and Google Sheets.
- These real-time tools have proven useful, helping the trainer or trainees coordinate their tasks or review someone's draft in real-time. Therefore, real-time editing is sometimes used when people draft articles and coordinate these real-time collaboration tools to publish them on Wikipedia.
- Many benefits of real-time collaboration were discussed:
    - Newcomers' familiarity with real-time tools.
    - The ability to exchange resources, expertise, and knowledge among trainees in real-time at the writing stage after training.
    - Roles emerge organically in real-time collaboration teams in that they involve distributing tasks among the group where each can focus on something they are skilled at.
    - The iterative process of reviewing in real-time, where more than one person contributes, ensures higher quality contributions.

Introducing real-time collaboration to Wikipedia

The challenges of introducing real-time collaboration were grouped into technical and social ones. Figure 6.12 and Figure 6.13, taken from the presented report, show the complexity of the

social challenges compared to technical ones. The challenges represented with circles and clouds symbolised ideas or a synthesis of different discussions about opportunities to address them and the research question.

6.5.2 Overcoming Challenges

As discussed in Chapter 2, Wikimedia's efforts in that area yielded good results. They have built an extension, a beta version of a visual editor with real-time collaboration functionality. However, adding a real-time collaboration feature affects many elements of how the community operates; due to many unanswered questions, this project stopped. Throughout the sessions, having an external tool was discussed to introduce an alternative model of collaboration with fewer social restrictions on an operational level. It should provide a space where people could have a tool to facilitate real-time collaborative editing outside of Wikipedia, with the possibility of incorporating it into Wikipedia in future.

From a technical perspective, countries that have a poor or restricted Internet connection may allow surfing webpages but not performing real-time editing. This technical challenge was discussed as a social one, adding social distractions in the editing environment that could disrupt real-time editing.

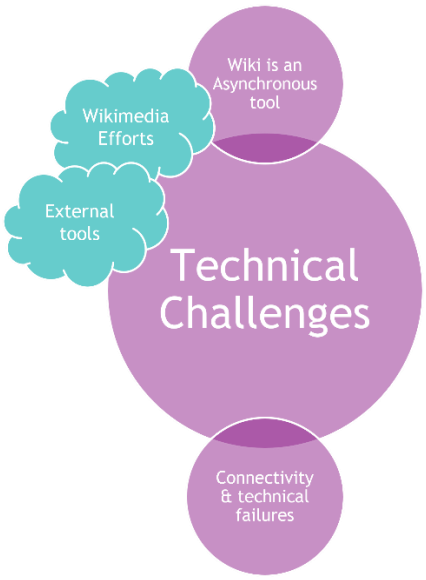


Figure 6.12 Technical challenges for introducing real-time collaboration to Wikipedia and possible solutions.

The community evolved around asynchronous editing, as did the rules, regulations, complementary tools, and extensions ecosystem. Figure 6.13 shows the main social challenges of bringing change to the norms discussed in the "Create" session, in which the thematic analysis results were shared.

Figure 6.13 Social challenges and possible solutions for introducing an alternative real-time collaboration technology to Wikipedia while at the same time being responsible and respectful of Wikipedia's rich social structure and history.

The possible solutions shown as clouds on top of the red circles are based on the thematic analysis. Many discussions pointed out that improving the mobile editing experience can solve many problems the community faces. However, since this is not directly related to real-time collaboration, it is included in the diagram but left without a direct connection to the challenges.

6.6 "Concept" Session – Insights and Findings

This session's first activity introduced a scenario where a group of staff members at a library interested in open knowledge are trained for real-time collaboration on Wikipedia articles. The participants reviewed, discussed, and evaluated the scenario of the training session before contributing to finalising the system concept (see 6.2.3).

Fostering Mutual Learning Through Participatory Design

The sessions generated insights for the research and facilitated learning. For example, participants benefited from what was shared from the literature and other participants' previous experiences, as discussed when Participant H learned about the PetScan tool (see 6.3). Another

example was a long thread (see Appendix D.9) that continued when Participant J asked the other participants about ways to solve the IP blocking issue faced by Nigerian newcomers after creating an account or making their first edit. Participants A and B provided multiple solutions and workarounds that were welcomed by Participant J, who said she would share them with her community.

This is a clear example of an issue that is unrelated directly to introducing a real-time collaboration tool. The Nigerian Wikipedia community already faces this social challenge in asynchronous collaborative editing. In the "Create" session, I made a quick decision: point the participant to where she could find my version of the answer in Figure 6.13 or step back and give an opportunity to another participant to address the question. I took the latter option for two reasons. First, this problem would affect real-time editing similarly, therefore, finding a solution by the community would contribute to this research question. Second, being an advocate for open knowledge contribution, I wanted to see if this session could solve a long-standing problem affecting those who consider themselves marginalised by the community of administrators. Section 6.7.2 discusses the different roles I played as a facilitator of the co-design sessions.

Fostering Collaborative Efforts to Address the Social Challenge of Real-Time Editing Attribution and Ownership.

At the "Concept" session, the discussion and opposing views demonstrated the complexity of the social challenges that the introduction of real-time collaboration poses. However, the advanced level of the discussions developed throughout the sessions helped the participants collaborate on filling the gaps in drawing the picture of introducing real-time collaboration through an external tool.

Figure 6.13 shows that all challenges have a possible solution cloud, but as shown in Figure 6.14, by the time these challenges had been presented to participants, I could not present a solution to one of the most complex social challenges: "Attribution and ownership". I was unable to identify a solution that could address this challenge through thematic analysis alone.

Figure 6.14 A screenshot from the Microsoft Teams call at the fourth session shows the social challenges slide before adding a solution to the attribution and ownership problem.

The missing solution led Participant F to take the initiative and wonder: "there is this social challenge that was identified, it is 'Attribution and Ownership', but the solution is absent [...] Could you further explain how you see attribution and ownership as a social challenge? Maybe that could also create suggestions or recommendations on averting this attribution and ownership as a social challenge."

I had a vague idea for an answer, but since this topic was one of the controversial topics about real-time collaboration, it led me to summarise the question about "Attribution and ownership" and provide several incomplete answers. One of the main questions is who submits and gets the credit for the real-time article formed outside Wikipedia. My proposition was considering whether Wikipedia rules allow posting other people's work under one person's name. Tracking the source of each added/removed character of an article formed using real-time collaborative writing is complicated but possible (Wang et al., 2015). However, communicating that to Wikipedia's API to register each edit to the account owner posed many social and technical challenges. Therefore, I collaborated with the participants to find out if it was plausible to add a brief description of the article's author identity within the summary required to be published when adding content on Wikipedia.

This caused Participant A to question the assumption that all seek attribution on Wikipedia. He added that the participants might all agree that one person would submit the

article and gets the credit. He believed that getting the credit for each person would be nice, and he continued saying: "I appreciate that does make online collab synchronous collaboration difficult if you're then trying to integrate it back into Wikipedia." However, Participant F disagreed, as she felt it was unfair or "biased" to attribute an edit to one person, excluding the rest of the contributors. This drove me to emphasise that the edit summary would give attributions to the rest. Another option could be asking each person to submit the section they contributed to the most, as Participant H does in his training sessions (see 6.3). Participant F was satisfied with this solution, however, before moving to the next session, Participant A flagged another possible danger in breaking one of Wikipedia's rules of preventing a "role account"[91] with one person submitting the created article under a role that represents a group. Then, I did not know what a "role account" was, which might disrupt the proposed system. To avoid wasting time, I continued the session, researched the topic, and reopened the discussion in the fifth "Create" session, as discussed in Section 7.3.

Universal Scenario for The Tool and Participant Owning the Scenario

The training scenario was designed to empower the trainers and trainees using a solution relevant to their challenges and ambition. As discussed in the previous chapter, it should have flexible components, enabling real-time collaborative writing and assigning different roles and tasks in a group.

However, going through user categories, personas, and the proposed usage scenario, Participant B questioned the universal applicability of said scenario. After explaining that being detailed in describing the group was a deliberate act of defining a workable solution, we could "put ourselves into their shoes". This method could be expanded on later rather than providing generic descriptions.

However, on the contrary, Participant E said,

> "It does seem to me quite universal that definitely fits with the kind of like profile of Edit-a-Thons that I've both led and attended as that you do get a real mix of people both with Wikipedia experience some good, some bad, some non-existent, and also that different levels of and areas of skills. So that you get some people who are very good at writing but haven't edited Wikipedia before or even are very good at history or whatever the subject is. And then I liked that you had a thing about one of them having a link, a

[91] https://en.wikipedia.org/wiki/Wikipedia:Sockpuppetry#ROLE

> personal family link to somebody that they tried to write about and then couldn't, because that to me also that kind of potential conflict of interest editing tends to come up quite a lot, especially if people have got some [...] so yeah, I basically my kind of first thought going through this was, firstly, it's super detailed and really like well done. I can imagine how useful it would be if you had all of that information about people when you started doing a training and then you could really tailor things [...] It struck me that I think, although obviously, you've got a specific group here deliberately, and they're from a, you know, a particular setting and stuff. And the different aspects of the kind of characters and profiles that you've picked out, I think. And probably are fairly universal, and in lots of ways, I would think based on ones I've done like I say, but I know it's a big world out there!"

Later, Participant J confirmed this by saying:

> "Previous experience is very, very relatable because according to Tala yeah, she said she has the assumption that Wikipedia is not very reliable, since anyone can edit, and Mary says she was surprised to learn that she can edit Wikipedia and amongst the other trainers and trainees. So, I want to say yeah, previous experience with Wikipedia is very relatable because before I started making edits on Wikipedia, I had no idea that I could [...] So, I want to say that this, this set of users that you have; they're very, very reliable. They're people who you can relate with, they are people, who I can relate with, because I've had previous experience as them and with all my experience throughout the past three years, I can say that it has changed a little bit."

This shows the efficiency of the collaborative persona method and the importance of infusing the individual personas' characters with a synthesis of the participant's previous experience. This led to more engagement and informed design decisions throughout the rest of the study.

Informed Collaborative Design Suggestions

Since the beginning of the sessions, participants have proposed many design considerations, suggestions, recommendations, and requirements. However, looking at the full spectrum, one can find that the further we go into the process, the deeper and more relevant and informed the design suggestions are. This maturity is identifiable through the type of questions asked by participants. A case in point is the question about the attribution challenge by Participant J.

As demonstrated above, Participant E linked her discussion about the collaborative persona to the ability to tailor the events to fit different skill levels. Later, Participant A would connect the discussion on the tool design to how this might spark change in Wikipedia:

> "I would still kind of encourage people to use the Wikipedia tools, but I think going back to the earlier discussion, we need to make it clearer how to get people unblocked [...] I know that a number of utilities and so on have done that on their websites that you can start a chat session or you can talk to a real person [...] I think having something where you can click on the button and start a discussion with somebody would advise on how to fix something would be a really useful thing [...] and maybe that's the sort of line that we should take Wikipedia down in terms of help have a button where you can join an online discussion about solving and answering queries about how to how to edit it and blocking IPs might well be a bit particular part of that."

Looking at the training journey maps, Participant F put herself in the "perspective of the trainee." She tried to represent someone who caused the article's deletion:

> "In trying to avoid such scenarios to keep reoccurring and to help motivate an editor who has created an article before, I think there should be [...] a step-by-step guide on how to maybe create each section of an article. Like I am seeing a scenario whereby you have someone trying to create an article [...] promotional, then you get a notification on that article that tells you why not use so and so words in place of this."

Participant J also suggested that a step-by-step guide on regulations would help trainees in our tool avoid getting their accounts blocked and/or articles deleted. Participant F confirmed that "there should be a section where new editors would get information about what they ought to do, what they ought not to do. I think it should also be included.". This prompted Participant A to say that such a feature could help ensure that the first sentence put context into why this article is important to the global audience, which provided reliable references. In Chapter 7, I outline how the participants' informed, collaborative design discussions based on the collaborative persona scenario helped shape the design significantly.

System Statement

The scenario helped identify and generate ideas for a system concept statement outlining its behaviour in order to deliver high value to the participants and appropriately consider the wider

community norms and socio-technical infrastructure. Given the little time to discuss the statement, participants were asked to comment in a follow-up email.

After reading the statement aloud, participants' comments were not directed at the statement, and the post-session follow-up comments also did not target the statement. However, the earlier insights were sufficient to improve the system concept and set the stage for the last phase of the PD process, where I transformed these concepts into interfaces designed in an iterative process.

6.7 Discussion and Summary

This section summarises the main findings and discussions shaping the next chapter.

6.7.1 System Design Decisions

Collaboration with participants and co-researchers led to the conclusion that introducing real-time collaboration is achievable and that it should provide the following main concepts and features:

- **External:** Provided through an external tool where people collaborate in real-time and move their contributions to Wikipedia.
- **Flexibility:** Real-time collaboration should be provided by a system built to provide Wikipedia trainers with all they need to tailor and deliver online training for trainees with diverse backgrounds. It offers an environment for groups of trainees to collaborate on learning, researching, and co-writing Wikipedia articles in real-time. If the trainee cannot join, it should have backup communication channels or tasks that accommodate joining from countries with a poor connection or limited access to a device.
- **Multifaceted Transparency:** The new system provides trainers with a flexible, transparent way to train the small groups on real-time collaborative Wikipedia writing, starting with the training process and ending with sustainable volunteer engagement. On the one hand, the tool should allow a high level of awareness for collaborators and the community, allowing trainees to be transparent with writing. On the other hand, the tool should facilitate the community's transparency with the reasoning behind their decisions after reviewing the article, especially if the decision involves deletion or IP blocking.
- **Support:** Trainees should be guided and able to request support during the writing process by the trainer or available experienced editors. There should be resources or tips on writing each article section in adhering to Wikipedia's community's rules.

Many design ideas and requirements had emerged at this stage of the research. These design concepts guide writing system requirements and filter through the ones mentioned already to help design the system interfaces driven by the sessions' findings. This contributes to addressing the second research question about the key features of a real-time collaborative tool for Wikipedia newcomers. The next chapter goes further in addressing this question by outlining these features and presenting a prototype that describes how they interrelate.

6.7.2 Participatory Design Process Implications

The phase discussed in this chapter has helped us position and narrow the design discussion down to specific yet generic users' needs. Providing the participants with fictitious usage scenarios close to their previous experiences helped significantly humanise their design insights about our users' personas journeys using the system. Returning to the research question **Q3 How can we co-design a tool and detailed training scenarios that would benefit from real-time collaborative editing?** Several characteristics of the ethnographically-informed distributed PD approach used in my research helped address this question. The section below covers four implications for the PD process.

Producing a Collaborative Persona Scenario

Researchers such as Grudin & Pruitt (2002) discuss the benefits of personas as a foundation for scenarios. However, they highlight the difficulty of getting the personas to suit the context. They argue that using personas "restores" elements, such as the prolonged participant engagement and considering the socio-political issues lost from contemporary PD practices. They attribute this loss to adapting PD to help with the current fast-paced, platform-independent software development needs. They note that the most reliable data to build the personas is founded on the observational or ethnographic study followed by user-led analyses.

When the scenario developed in this PhD research matched the participants' experiences, some advocated for certain design decisions or finding solutions to issues, such as IP blocking. This also made those who could not relate at least sympathise with this need, clarifying misconceptions about internal bias or revealing assumptions and addressing them with a constructive discussion and design decisions that can mitigate the problem. The diverse social issues that emerged from discussing real-time collaborative editing elaborated on the importance of maintaining the social equilibrium between the researcher and informed participants. Researchers should use methodological approaches that empower the community participants to direct their platform's evolution with usage scenarios. Such an approach

preserves connections that bring a global community, such as Wikipedia's, together to create an international encyclopaedia.

Playing a Pragmatic Role as a Researcher

Reaching such a stage of identifying themes for answering this research question while being the only organiser and facilitator of these events meant wearing several hats. This was an important element in addressing the Conceptualisation phase research question.

Most research, such as Grudin & Pruitt's (2002), describe PD as a resource-intensive process. Running these sessions requires thorough planning by organisers, sustained communication with different stakeholders, a venue, and much more. However, I would argue that by providing multiple options for methods combined with the pragmatic role the researcher can play, there could be a way to preserve the values of PD with available resources. Being pragmatic in my approach, I was able to conduct the study on my own. This necessitated awareness in situations that demanded specific strategies to drive the PD process forward.

Dahl & Sharma's (2022) analysis of 14 interviews with PD practitioners led to the following categorisation and facets of the PD facilitator's role:

- **Trust builder.** Building a safe space for the participants, leaving assumptions aside and exchanging ideas.
- **Enabler.** Defining a "common language" and means or tools to support the participants in expressing their thoughts.
- **Inquirer.** Understanding the needs through well-designed questions (open/semi-open) and getting to the roots of the issues through follow-up questions.
- **Direction.** Directing the discussion back to the research question without disrupting the flow of thoughts. When the debate crossed the study boundaries, I pointed it out, helping me identify insights and connections.
- **Value provider.** This was evident with insights, realisation, and opportunities to learn from the contributions, such as new tools and practices that increase productivity.
- **Users' advocate.** Depending on the stage of the research and discussion, I advocated for specific users in designing decision-making. Wikipedia is unique in that the community, especially veteran editors, hold immense power over newcomers, and lobbying for them could be challenging.

I would add another facet to Dahl & Sharma's (2022) list, the **Observer**. At certain stages during the session, I stepped back from my role as a researcher to observe the conversation

among participants. This included holding back on answers to the participants' questions, knowing that my intervention would disrupt a conversation flow that might result in an interesting insight. This was apparent in breakout rooms, where one participant decided to take the leader of the group's role and go through the assigned tasks. This led to taking ownership of the innovation process and finding solutions, higher levels of mutual learning among the participants, and questioning their assumptions by having an open forum to find a working scenario for the group.

Tools for Sharing the Decision Power with Newcomers

Participatory Design is well known as the go-to method for empowering the end-users of a system in a highly politicised environment where users' voices are barely heard (Beck, 2002; Shapiro, 2005; Chisik & Mancini, 2019). However, researchers, such as Beck, claim there is a gradual decrease in focus on politics in the PD researcher's agenda. The author argues against arguments that diminish the PD role at an age where "political concerns remain on the minds of many" (Beck, 2002). This was evident in a study that reviewed publications at the PD conference and found that a few articles had such a political element. (Bergvall-Kåreborn & Ståhlbrost, 2008).

As discussed in Section 4.3.7, the power dynamics that govern Wikipedia are unique. In what follows, I argue that the time is opportune for a conscious review of the Wikipedia community's rules, policies, and power dynamics. Change is necessary for how the Wikimedia Foundation and its core community use their power in design decisions.

Many have researched the dedication of "elite" editors leading to Wikipedia's ubiquity (Kittur et al., 2006). With such recognition of these few editors comes power and higher ranks in the "core community". On the other hand, a bulk body of knowledge exists on the marginalisation of certain groups based on gender (Wagner et al., 2016; Ford & Wajcman, 2017) or geographical location (Graham et al., 2014; Graham & Hogan, 2014; Ford et al., 2015; Ford, 2017). Ford discusses her ethnographic study with a small group of Kenyan Wikipedians who describe "their hopes and disappointments regarding the community and the platform." She argues that most research focuses on "those who are highly active and visible on Wikipedia" and that there is a necessity to "study the lacunae, the people who remain on the edges of Wikipedia because of disappointment, rejection or invisibility" which backs the arguments made by the Nigerian participants in my research (see 6.4.1). Ford encourages researchers to use an ethnographical approach that helps them "experience the position of the uninitiated, the newcomer, the powerless within the network." (Ford, 2017). Ford & Graham

(2015) give examples of clear negative consequences on society as an effect of digital exclusion, limiting the marginalised community's ability to represent themselves online in a world "increasingly governed by the logics of the semantic web."

Bjork-James, (2021) argues that changing the demographics of the editors and writing about the marginalised community is inadequate to fill knowledge gaps and overcome the "systemic biases". The author notes that observers and numerous Wikipedia editors "have argued that the systemic biases of Wikipedia are inherent to the current global distribution of knowledge production, and can only be overcome by changing the encyclopaedia's standards of inclusion" (Bjork-James, 2021, p. 207).

Recent stats, as shown in Figure 3.2, indicate that most Internet newcomers come from developing countries. The number of people accessing the Internet from the Middle East and North Africa has almost tripled during the last decade, while for North America, it has only increased by almost 20%. As discussed in Chapter 8, this PD process benefited from Hagen et al.'s (2012) framework, contributing by offering the Wikimedia Foundation to revise its inclusion standards and help the community have a safe online space and fruitful discussions welcoming newcomers. This process is equipped with several methods that ensure informed participation in the innovation process, bringing more in-depth insights for change that extend the limited available types of engagement. This should subsidise other community decision-making mechanisms, leading to adding, rejecting, or removing features with complex socio-technical ramifications. This will also help share the decision-making power with newcomers who cannot navigate the software development process while at the same time sustaining engaging in daunting voting debates.

Managing The Expectation of Thematic Analysis as a Tool for Co-design

Conducting the thematic analysis emerged as a necessity to make sense of the bulk-generated qualitative data. The aim was to identify solutions or design ideas for most challenges. However, it was evident that the TA alone was insufficient to untangle the complexity that emerged in the Conceptualisation phase and that further discussions were needed to complement the analysis findings. As will be discussed in Chapter 7, this has led to the use of Activity Theory as an analytical lens to explore the accumulated data from the observation study and co-design sessions to list requirements for the WikiSync system.

6.7.3 Summary

This chapter explores how experienced and newly trained Wikipedians envision benefiting from real-time collaboration considering the specific team characters and their journeys using WikiSync to achieve their goals.

It described the set-up and findings from the “Position”, “Concept”, and two “Review” sessions that were conducted as part of this Conceptualisation phase. In these sessions, participants were asked to collaborate on formulating the opportunities to address these challenges by co-designing personas and training scenarios using real-time coordination and editing tasks. The chapter then discusses the “Concept” session, where these personas and scenarios are reviewed and used as a foundation for proposing early-stage journey maps and a conceptual tool. The next chapter translates the output of this chapter into design sketches.

Chapter 7 Design and Social Voting Phases

As discussed in previous chapters, I have divided Hagen et al.'s (2012) framework into three phases. In Chapter 6, I have described the Conceptualisation phase (including "position", "concept", and "review" sessions) that resulted in defining the concept behind the WikiSync system. This chapter covers the ***Design*** phase of how design concepts were analysed to develop the first sketches of WikiSync, then covers the iterating process leading to the final prototypes.

Since the ***Social Voting*** phase is a form of consulting and iterating on Design phase findings, Chapter 7 covers both phases, leading to the prototyping of WikiSync –the first Wikipedia training tool for designing an event that involves real-time collaborative editing of Wikipedia articles.

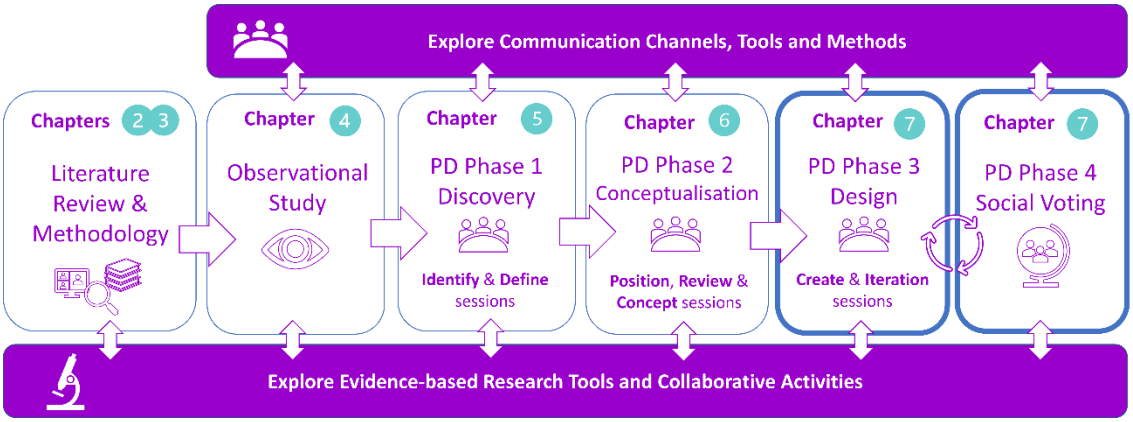


Figure 7.1 The last phases of the participatory design process are highlighted in relation to other stages of this research study.

The aim of the Design and Social Voting phases is to answer this research's **Q4 How can we apply and possibly adapt participatory design processes for online communities and Wikipedia specifically?** through addressing the following three questions:

Q4.1. How to untangle the socio-technical complexity of introducing the new collaborative writing process?

Q4.2. How to connect the system requirements with the qualitative results of a thematic analysis?

Q4.3. How and when to scale up and down participation in designing a tool that serves a large and widely distributed community?

Answering the above questions resulted in a prototype for WikiSync tailored to the community's needs, leading to high acceptance among the participants regarding the

applicability of such a collaboration model in Wikipedia training. It also has helped identify an innovation process for introducing change in Wikipedia.

As shown in Figure 7.2, the Design phase covers an Activity Theory analysis to process the Conceptualisation phase's outcome and produce user requirements which form the foundations for sketching the system on Miro Board for the "Create" session.

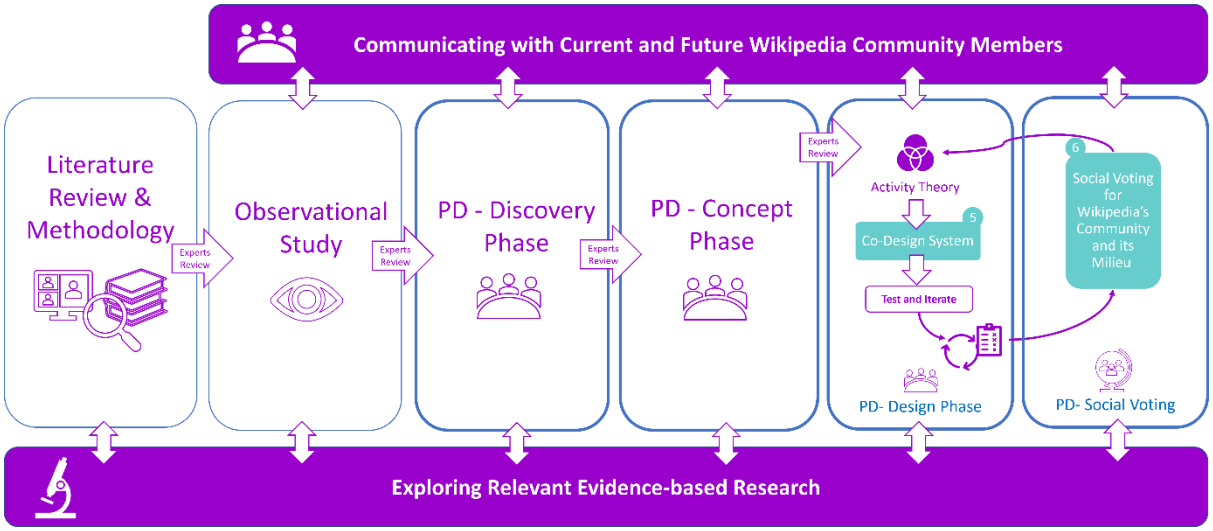


Figure 7.2 The details of the last phases of the participatory design process.

In the Design phase, participants reviewed the final system concept and sketched new designs for the system interfaces in the "Create" session. The feedback was incorporated into the design, and participants were invited to an iteration of the "Create" session for another review round focusing on pitching ideas to improve the real-time collaborative space interface through sketching. These two sessions have led to deciding how the system should look and behave and to planning to consult the broader community on our findings.

The Social Voting phase starts with reviewing the design ideas, which results in the Design phase with participants and experts, iterating and translating the output into a list of ideas that opened the discussion for the broader community allowing the design process to scale up and benefit from the "wisdom of the crowd". Next, the community's and research collaborators' feedback is analysed to iterate on the design at multiple points. Finally, a high-fidelity design is developed using Figma[92] to be shared with the community for a final evaluation session.

The Design and Social Voting phases rely on similar software, pre-session, and introductory activities discussed in previous chapters, with variations in tools or methods used

[92] https://www.figma.com/

to facilitate activities designed to achieve the aim of the session in focus. For example, the ***Social Voting*** phase benefits from online social ideation and a voting tool called Tricider.[93]

As shown in Figure 7.2, Section 7.1 covers the Design phase's activity theory analysis. Section 7.2 will briefly cover the iterative process that has led to the low-fidelity sketching of the WikiSync system. Finally, Section 7.3 covers the co-design cycles that expand the participation to the broader community.

7.1 Activity Theory as an Analytical Lens

This section covers the use of Activity Theory (AT) in my research. After a brief history and evolution in Sections 7.1.1, Section 7.1.3 discusses this analytical framework's capabilities in supporting the design of Information Systems (IS). The logic behind using AT in this research context is highlighted by sharing cases that yielded a successful outcome. After discussing the principles of AT in Section 7.1.4, Sections 7.1.5 to 7.1.8 discuss using AT to untangle the socio-technical complexity of introducing the new collaborative writing process and connect the qualitative results of a thematic analysis with the insights and requirements for designing the system.

As covered in Chapter 6, the Conceptualisation phase discussions about challenges and tensions spiralled out of the assumed scope of research (real-time collaboration model). However, there were several clues that these discussions had direct or indirect connections to real-time collaboration. For example, mobile compatibility is a significant challenge for Wikipedia's asynchronous editors and will affect real-time collaborators too. In addition, there are several other examples of various challenges with different proximity levels to real-time collaboration. At this stage, the need exists for a research tool to help clear some of this complexity before sketching the system, and the Activity Theory emerged as a viable framework in my research context.

Activity Theory is a meta-theory, rather than a predictive theory, which helps understand the "everyday practice in the real world [. It] is a powerful and clarifying descriptive tool rather than a strongly predictive theory." (Nardi, 1995, p. 4). Unlike Thematic Analysis, which was used to narrow the focus on the emerging themes of my research project, AT was to explore, define, and analyse the broader picture of introducing real-time collaboration. This has provided the needed connections, narrowing the scope of my research again on real-time collaboration. Through analysing several activity systems, the framework is used as an analytical lens for data from sessions to produce culturally fit requirements for WikiSync.

[93] https://www.tricider.com/features

Since its inception, the large body of literature around AT has led to prolonged, constructive debates that have led to its evolution through several phases and generations (Engeström, 1987). However, this has engendered many views on how, when and where it should be employed. Therefore, the publications on AT use cases might differ based on the maturity of the activity theory used at that time and the research's unit of analysis and focus.

Therefore, special care has been given to exploring relevant literature and choosing what fits this study's theoretical foundations. However, there has been a need to strike a balance between the time needed for a cautious approach to using the AT and acting quickly and pragmatically to prevent delaying the next promised "Create" session for participants.

Scope of Use

Leont'ev defines ***actions*** and ***operations*** as the underpinning of an ***activity*** (Leont′ev, 1978). The author describes how an ***activity***, ***actions*** and ***operations*** are linked in a hierarchical structure, where the ***activity*** consists of a chain of actions and operations. An ***activity*** is generated by the motive, which, in turn, determines the subjects' (or actors in an activity) specific goals. These actions are composed of operations determined by the conditions in which the activity is performed. As shown in Figure 7.3, Kuutti (1995) illustrates hierarchical levels of activity with some examples.

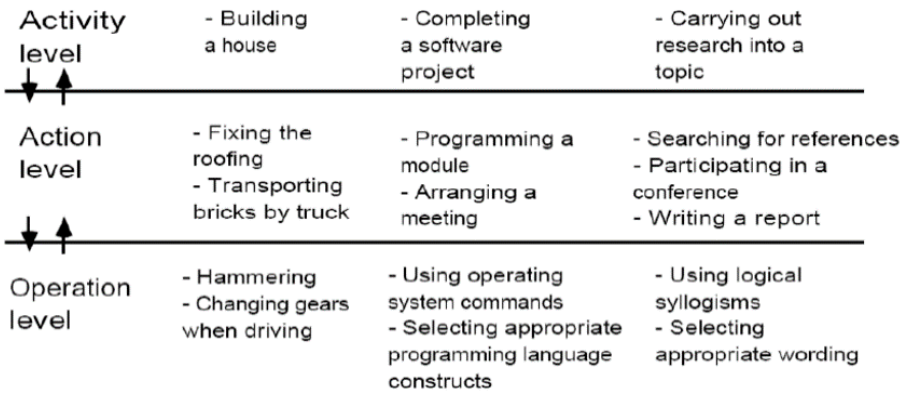


Figure 7.3 An example of the hierarchical levels of activity (Kuutti, 1995, p. 33).

Gonçalves et al. (2013, p. 554) discuss the type of questions for these levels that correspond to different levels of answers: the ***activity*** level focuses on "***why*** things happen and is developed over a long period of time within a socio-historical process", whereas the ***actions*** and ***procedures*** focus on the what and ***how***. For example, Hautasaari (2013) uses this hierarchical level of analysis to answer the ***how*** type of questions about the Wikipedia editors' ***actions*** in translating a Wikipedia article and found three distinctive strategies non-expert translators use

to translate articles from English to Chinese. However, this PhD research focuses on the macro level by studying ***activities*** as a network of activities without diving deeper into the ***actions*** and ***procedures***. An analysis of the historical development of some of the activities (such as Wikipedia training and asynchronous collaboration), their components, and some tension that has evolved through time will be provided in this chapter, leading to design insights and requirements.

Section Outline

The use of the Activity Theory (AT) as a theoretical framework is structured as follows: Section 7.1.1 looks at the AT evolution. Section 7.1.2 covers the reasons why this project uses AT. Section 7.1.3 explores literature that helps define some aspects of the framework relevant to my research context, such as the uses of AT in collaborative tools design, and later concentrates on Wikipedia cases. Section 7.1.4 covers Engeström's (2001) work, summarising the activity theory's five principles and interpreting them for my research context. Sections 7.1.5 to 7.1.7 focus on the fourth principle and the role of contradiction within and between activities and their components (subject, tool, object, community, rules, division of labour). Section 7.1.8 covers the listed contradictions in conjunction with the socio-technical challenges from research data in earlier phases to map a new Wikipedia training activity in a network with other real-time collaborative writing and revising activity systems. The discussion shows how looking at this network of activities, including my research's PD activity, has helped bring more clarity and elaboration in the activities' ecosystem to devise the system's conceptual model and requirements that could be tracked to the challenges from the data.

7.1.1 Activity Theory Evolution

The Activity Theory is a well-researched framework, with plenty of publications on projects using it in many contexts and at different levels and areas of research. Therefore, after defining it, the exploration phase focuses on unpacking the framework's history and constant development to identify similar studies and define its utilisation scope for this PhD research.

Activity Theory is defined as a "cross-disciplinary framework for studying different forms of human practices as development processes, with both individual and social levels interlinked at the same time." (Kuutti, 1995, p. 25). Kuutti (1991) describes AT as one of the few social sciences schools of thought to originate in the Soviet Union and be adopted in the Western World as a framework for analysing human activity on a social and individual level. It was developed by researchers from the field of social sciences, and later expanded to be utilised by a multidisciplinary, international audience. Notably, Vygotsky 1920s, L. S.,

Leont'ev, A. N. and Luria, A. R. are considered the early originators of the concept that most contemporary research referred to as the Cultural-Historical Activity Theory (CHAT) (Barab et al., 2004).

Activity Theory identifies human activities in their social context as the minimal unit of analysis (Kuutti, 1995). However, it was previously considered the human action, which is relatively easier to examine in design laboratories without the context but "less fruitful" for real-life situations (Kuutti, 1995). Including the context, according to the author, makes activities' object a collective endeavour, even when the focus is on specific actions.

According to Vygotsky, the basic form of all human activities has a third element of artefacts or instruments that mediate between the activity performer and its object, as shown in Figure 7.4. Mediation could take a positive or negative form, as Kuutti explains:

> "The tool at the same time both enabling and limiting: it empowers the subject in the transformation process with the historically collected experience and skill ``crystalised'' to it, but it also restricts the interaction to be from the perspective of that particular tool or instrument only; other potential features of an object remain ``invisible'' to the subject" (Kuutti, 1995, p. 27).

This connects to insights from analysing the discussion in Chapters 5 and 6, where the core Wikipedia community members view the Wikipedia article as an object from the perspective of asynchronous editing tools with some features, such as the ability to collaborate on it in real-time "invisible".

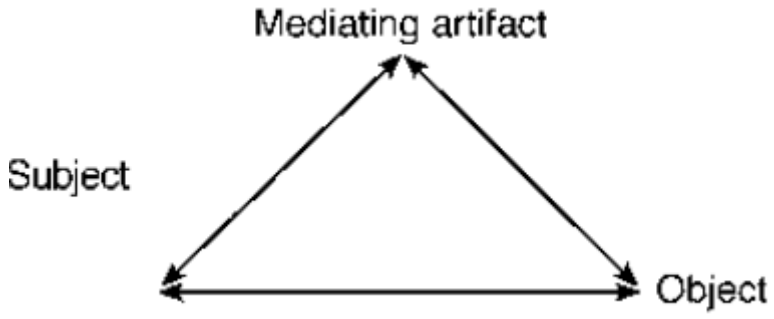


Figure 7.4 An adapted version of Vygotsky's basic scheme of a mediated activity (Engeström, 2001).

Activity Theory Generations

Each human activity is mediated to achieve change in the results. However, Vygotsky's first-generation model in Figure 7.4 is "too simple to fulfill the needs of a consideration of the

systemic relations between an individual and his/her environment in an activity" (Kuutti, 1995, p. 27), which led Engeström to add the community as an additional main component to the model, which shares the object with the subject, (Kuutti, 1991). As shown in Figure 7.5, the updated activity system model proposed by Engeström (1987) represents the rules that mediate the relationship between community and subjects, while the division of labour mediates the relation between object and community (Engeström, 2014). Activities are recognised and distinguished by the purpose of the activity or object (for example, solving a problem or writing a Wikipedia article), and transforming this object into an outcome stimulates the activity to exist (Kuutti, 1995). Outcomes could be the knowledge formed from experiencing the activity, such as learning from writing a Wikipedia article.

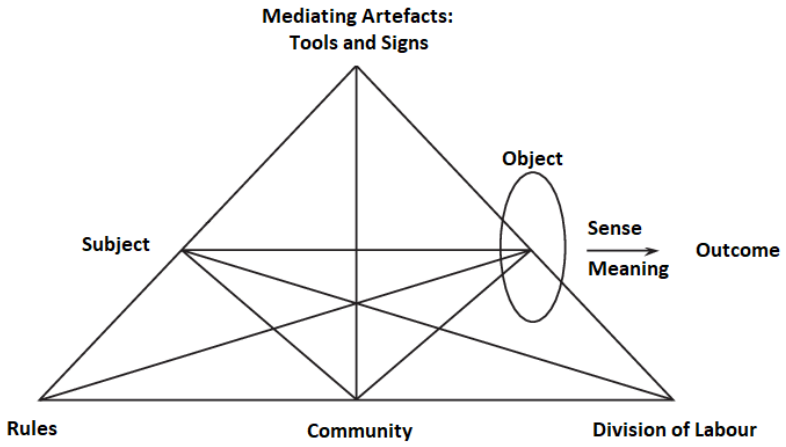


Figure 7.5 The structure of a human activity system (Engeström, 1987, p. 78).

Activity Theory considers the contextual aspects of the historical development of a phenomenon over time, whether it is a new technology or a learning method used by those involved in the activity system (Jonassen & Rohrer-Murphy, 1999). One common use of Activity Theory can be seen in studies, such as Bryant et al.'s (2005) (Becoming Wikipedian) and Hewitt's (2004) (community exploration) as a sense-making tool. It systematically reviews the following six interdependent elements of a socio-technical activity system using the diagram shown in Figure 7.5.

Below are the Activity System (AS) components, adapted from Kuuti (1995):

1) Subject(s) could be a group of people or an individual (sometimes called actors), that undertake an activity to transform the object into an outcome.

2) The object(s) could be something tangible, as a document or lesser material, such as a process, or nonmaterialistic, as an idea with the intention to transform by subject(s) into an outcome.
3) Tools (or mediating artefacts or instruments) support subjects' actions to interact with the object and hold into it the "historical development of the relationship" between subject and object. It could be software, sign language, or guidelines used to describe or explain the objects.
4) The Community component covers the social and contextual aspects of the activity. The community that shares the object conveys the rules and social norms for the activity. It could represent the team members performing a task or a subject's social environment, which could influence their behaviour or engagement with the activity; for example, the Wikipedia community.
5) Rules and Norms establish and dictate the relationship between the subject and its community. For instance, Wikipedia community policies, guidelines and procedural information.[94]
6) Division of Labour is the distribution of activity task responsibilities. Activities usually involve the contributions of many individuals who collectively make the activity possible. For example, writing an article by one individual on Wikipedia involves the support of Wikimedia and its community of web developers and interface designers to provide the technology needed to perform the activity.

7.1.2 The Rationale for Using the Activity Theory

Chapters 4 to 6 covered the reasons why real-time collaboration is missing and the opportunities to introduce it in Wikipedia training. This has many aspects, including Wikipedia's current innovation and design process. Understanding how this can come together in shaping WikiSync design has been necessary. For example, the Wikipedia community and my research participants discussed real-time and asynchronous editing co-existence as one of the main reasons leading to the halt of real-time collaboration. How to avoid an edit conflict between those editing in training with real-time collaborative editing and those editing asynchronously in the current established community? This question, in addition to many other concerns where socio-technical complex issues are hard to explore with conventional analytical tools.

[94] https://en.wikipedia.org/wiki/Wikipedia:List_of_policies_and_guidelines

The goal is to address concerns about introducing real-time collaboration without affecting the Wikipedia community environment, as addressing one PD participant's concern may contradict one of the community rules or available tools. Bringing a solution to one PD participant in my research could contradict another PD participant's need to address a different concern on the same topic (as seen in the IP blocking discussion in Chapter 6). Another factor concerning the complexity of these contradictions is the long history behind the development of community standard operation procedures, rules, quality control, usage of tools, and the way they distribute work.

Therefore, any trajectory for a real-time collaboration process should be guided by why a certain change is needed and informed by the different systems and components that interplay in forming the historical culture of the community.

Thus, it is evident that the complexity of my research's social and technical challenges needed a theoretical lens to untangle it. Devising design solutions to change community practices is a complicated endeavour requiring considering many possible tensions (Barab et al., 2004).

> "The goal of improving the world is a messy business, with numerous struggles, opposing agendas, multiple interpretations, and even unintended and controversial consequences. Instead of simply building an artifact to help someone accomplish a specific task, the goal is to develop a design that can actually support the user (and the culture) in his or her own transformation." (Barab et al., 2004, p. 210)

In this quest for analytical tools, AT surfaced as a potential sense-making framework to analyse the complexity of the relations between Wikipedia's training and collaborative writing activities. Wikipedia involves several technical factors, such as Wiki technology, and some social ones, such as the cultural norms developed by the editorial community since its inception in 2001(Müller-Birn et al., 2013). Providing an alternative collaboration model for knowledge construction must consider the historical development of those factors. Therefore, choosing AT was driven by the importance of acknowledging the community history when co-designing a solution to introduce real-time collaboration. This ties back to my PhD research's aim to bring culturally aware solutions through collaborating with the community.

Sections 7.1.5 to 7.1.8 demonstrate how AT has afforded an opportunity to organise, project, and interpret some of the data collected in my research through several methods,

driving forward the discussion on the proposed real-time collaborative Wikipedia training scenario.

A brief history of AT for design is covered in the following section, leading to use cases in generating Information System (IS) requirements.

7.1.3 Activity Theory for Design

Activity theory has been used to analyse activities in many fields, including as a theoretical tool in designing courses or learning environments and modelling collective knowledge designs and conceptualising the use of digital technology for writing (Jonassen & Rohrer-Murphy, 1999; Hasan, 2003; Engeström, 2014; Abdullah, 2014; Said et al., 2014; Johnson, 2016; McCalla, 2019; Al-Ali, 2020; Augustsson, 2021).

Researchers have recognised the activity theory role in innovation studies, HCI and CSCW as a suitable framework for deriving insights out of analysing the transforming activities in the socio-technical infrastructure of computer-supported systems (Bødker, 1990; Kuutti, 1991, 1995; Hyysalo, 2000; Miettinen & Hasu, 2002; Bertelsen & Bødker, 2003; Bødker & Klokmose, 2011, 2012; Good & Omisade, 2019). Bertelsen & Bødker (2003, p. 249) describe AT use in HCI as "a set of conceptual tools, rather than a collection of tools and techniques ready for practical application". AT provides a theoretical foundation for informed insight to design and assess computer software, focusing on the historical development of practices and the context in which a system has been used. In addition, exploring the benefits of using AT in IS development is evident in the publications that used it to guide the research questions and the design research. For example, Mwanza (2001) and Al-Ali (2020) used AT to influence the questions being asked of the participants for designing solutions. Moreover, Good & Omisade (2019) describe the link between AT throughout the stages of the UCD cycle in the design and evaluation of two healthcare systems, emphasising the role AT can play as a theoretical lens in UCD. Researchers have used AT in different ways and stages as a lens to examine the data and identify connections between the various elements of the activity system to analyse socio-technical systems, resulting in design decisions and requirement elicitation (Turner & Turner, 2001; Collins et al., 2002; Uden, 2007; Georg, 2011; Ashritha et al., 2017; Dimitrakopoulos & Uden, 2020; Durst et al., 2020). Researchers, such as Hyysalo (2000, 2005) and Abdullah (2014) have used AT to untangle complexities, identify design work tensions, and suggest alternative work models.

In this PhD research, collaboration activities (writing articles and onboarding) are examined to identify the historical development of tension, culminating in design insights and

needs. Since the emphasis is on training and collaborative writing activities, the PD activity discussed in Chapters 5 to 6 is not analysed and will only be mentioned later in this chapter in relation to its object (a training scenario that uses WikiSync), which is used as another activity's mediator.

AT for Analysing Peer-Knowledge Production Activities

Several research papers discuss the use of AT to analyse peer-knowledge production activities. For example, Hewitt (2004) has done an extensive activity system analysis to understand the complex nature of face-to-face and online interactions to improve learning and their virtual knowledge-building community of practice. This detailed analysis and the wider adoption of AT have influenced several researchers to study Wikipedia using this framework (cf. (Bryant et al., 2005; Slattery, 2009; Petrucco, 2010; Hautasaari, 2013; F. Liu et al., 2020). Using AT, the researchers contribute to advancing the understanding of Wikipedia as discussed in the following research cases:

Bryant et al. (2005) use the framework to describe their research study results which use another research framework -the legitimate peripheral participation.[95] Their results explain the factors that facilitate the transitioning activity of the newcomers into active editors. Through AT, Bryant et al. are able to describe the socio-technical factors of becoming an active editor and how their perception of Wikipedia and the need for (and use of) its features change over time. Initially, their research participants expressed their lack of knowledge of the community and the division of labour behind the articles. They assume that "Wikipedia seems more like a collection of articles with random people adding information here and there than like a collection of people talking about, editing, and protecting their efforts to author good work." (Bryant et al., 2005, p. 7). They do not mention reading the activity history or engaging in the articles' talk pages; however, these tools have become more useful and needed as their role, engagement, and level of support from the community change over time. They also note that "Wikipedia supports a more robust set of activities for Wikipedians than for novices." (Bryant et al., 2005, p. 6).

Hautasaari (2013) uses the Activity Theory to drive design decisions for a system that supports the articles translation activity in Wikipedia.

Slattery's (2009) work utilises the Activity Theory in the context of Wikipedia's fact-building activity to demonstrate the importance of the socio-technical link between technology

[95] In a community of practise or collaborative project, legitimate peripheral participation (LPP) defines how beginners develop into seasoned participants and then turn into project veterans. (Lave & Wenger, 1991)

and fact-building communities. The author highlights the importance of understanding the activity before designing functionalities that support the system.

Petrucco (2010) uses AT to look at a Wikipedia article-writing activity that uses Wikipedia as a tool for knowledge construction and learning outcomes.

7.1.4 Activity Theory Principles

According to Engeström (2001), the five principles are as follows:

The first principle is that an activity system having an object mediated by artefacts in its network of activity systems is considered a single unit of analysis. The different components of an activity system and its goal-oriented actions and operations could be independent but only understood in the context of the entirety of the system as subordinate units of analysis. For example, ongoing vandalism protection activity is in the network of and connected to Wikipedia's newcomer onboarding activity.

The second principle discusses the "multi-voicedness" of community views, norms, and traditions of an activity system. The division of labour in the activity system leads to differences in interpreting an activity system. This could form challenges that lead to learning processes and innovations in addressing them.

This is apparent in how different community members interpret the reason for deleting an article in the Wikipedia community, which could be regarded as a necessity to increase the overall quality of Wikipedia, but at the same time, it is interpreted as bias against newcomers by others.

The third principle is that the activity system evolves through time. Therefore, the problems should be understood in their historical context of development.

For example, considering the historical development of the editing activity in Wikipedia would help me understand why the current object of the activity is an asynchronous article, rather than a real-time one. In the same way, introducing real-time collaboration when Wikipedia launched would have been different than presenting it two decades later to the community. The Internet, as used by online communities this decade, is shaped by the ever-evolving social media platforms and the increased reliance on free real-time web applications that serve users' productivity.

The fourth principle emphasises the role of contradictions as a driver for change and development, which this thesis covers in detail in Sections 7.1.5, 7.1.6, and 7.1.7.

The fifth principle describes the activity system development as a cycle that starts with individuals questioning contradictions and diverting from the assumptions or long-lasting

traditional way of conducting an activity. This transformation expands, affecting other activities in its network to encompass a wide range of possibilities for change and improvements. This may involve transforming and re-conceptualisation of an activity object and motive in a way that "embrace a radically wider horizon of possibilities than in the previous mode of the activity." (Engeström, 2001, p. 137). This connects to Section 7.1.8, where the new real-time collaboration activity system is illustrated to explore how it could interact with Wikipedia's traditional asynchronous collaboration, and then place it as part of a more comprehensive network of activities in Wikipedia training.

7.1.5 Contradictions for Driving Development and Change

This section discusses contradictions as a concept for driving development and a wider scale of change leading to the solution. Not to be confused with problems or conflicts, in an Activity System, contradictions result from historical tensions inside or between activity elements or more than one activity system. Ilyenkov first identified contradictions in 1974 as the "historically accumulated dynamic tension between opposing forces in an activity system." (Karanasios et al., 2017, p. 2). This concept has evolved with the new Activity Theory model brought by Engeström in 1987, who considers that "contradictions of activity remained an extremely touchy issue" (p. 6) and a "driving force" for providing insights and a practical guide to identifying solutions that bring organisational change and transformation to an activity. Karanasios et al., 2017 discuss that there is interchangeability in using the term "contradiction" with tension in research to describe a historically emergent phenomenon. My thesis does not go into the debate over the terminology and relies mainly on Engeström's interpretations of this concept.

Engeström (2014) builds on the following words of Ilyenkov to explain the nature of this phenomenon that is capable of producing solutions that emerge from the tensions between the component of an activity and its surroundings. Section 7.1.8 attempts to link this to the change that my research aspires to in addressing the introduction of real-time collaboration in Wikipedia training and possibly influencing editing in Wikipedia:

> "In reality it always happens that a phenomenon which later becomes universal originally emerges as an individual, particular, specific phenomenon, as an exception from the rule. It cannot actually emerge in another way. Otherwise, history would have a rather mysterious form.
>
> Thus, any new improvement of labour, every new mode of man's action in production, before becoming generally accepted and recognised, first emerge

> as a certain deviation from previously accepted and codified norms. Having emerged as an individual exception from the rule in the labour of one or several men, the new form is then taken over by others, becoming in time a new universal norm. If the new norm did not originally appear in this exact manner, it would never become a really universal form, but would exist merely in fantasy, in wishful thinking." (Ilyenkov, 1982, p. 83)

Bryant et al. (2005) discuss the resemblance between the concept of producing solutions that adapt to the needs and the early Wikipedia founders' "brilliance" story of converting the traditional authoring model and adapting it to the needs of volunteers for an easy editing process. However, I would argue that this adaptability, as well as the innovation activities, have become rigid, forming one of Wikipedia's challenges over time. As discussed in Chapters 1 and 2, making changes to adapt to the newcomers' needs is a complicated process for the Wikipedia community, especially their veteran editors. This leads to attempting to work around this resistance to change, as seen in the use of real-time collaboration tools, which takes place, but has not been adopted by Wikipedia trainers or the community due to several social and technical challenges.

Sections 7.1.5 and 7.1.6 utilise Engeström's method of contradiction classification and cover examples of the causation by analysing the tensions inside the activity of producing articles using real-time collaboration tools and the tension between the training activity and other activities that depend on asynchronous collaboration. First, the following section describes the levels of contradictions.

Levels of contradictions

Before covering the contradictions in the activity systems under study in my research, looking at an example of different levels of contradictions from Foot's (2014) research based on Figure 7.6 would help one understand the direction my research took in identifying the contradictions. The list is guided by the four levels of contradictions explained by Engeström, 2014. Each level is followed by concise examples from Foot's (2014) research to help frame my research's contradictions later. Foot uses the CHAT framework to analyse the tension in relation to health clinics activity in the United States that involves doctors, patients, and professionals:

Primary contradictions happen within the element node itself. Engeström discusses how capitalist economy commodities, for example, have the dual nature of being both a use and an exchange value, leading to possible tensions or contradictions.

For example, (Foot, 2014) explains that a clinical activity could have a dual nature object of nurturing health for patients (use value) and making a profit (exchange value) to sustain its costly operations through financial gains. Primary contradictions are ordinary and exist in all nodes at all times.

Secondary contradictions emerge when the tension caused by the primary contradictions escalates.

Following the same example, an insurance company can impose rules on the clinic's doctors to only view patients for 15 minutes to increase profit. This may bring about an increase in tension for doctors who have the dual motive of making a living and treating their patients at once. This secondary contradiction lies in the relation between the rules and the object that has evolved from the dual nature of the activity object.

Tertiary contradiction takes place when a more "culturally advanced" activity system's object is proposed into the system to address or alleviate one or more than one secondary escalated tension.

Following the same clinic example, a nurse hired to assess the patient's situation for getting an appointment could suggest a more equitable approach (activity system) to maintain their services efficiently without relying on the insurance company. This would "precipitate" tertiary tension, with some people accepting her suggestion and others opposing it, with the probability of triggering a motion for change, redefinition and development phase of the central activity, depending on the power relations of the division of labour.

Quaternary contradictions are tensions between an activity system and other activity systems that access or contribute to its activity.

Again, going back to the clinic example, if the staff decided to adopt the changes of the object by turning the clinic into a non-profit, it would stimulate a knock-on effect, leading to contradictions or tensions with the insurance company activities that influence the clinic activity system's rules.

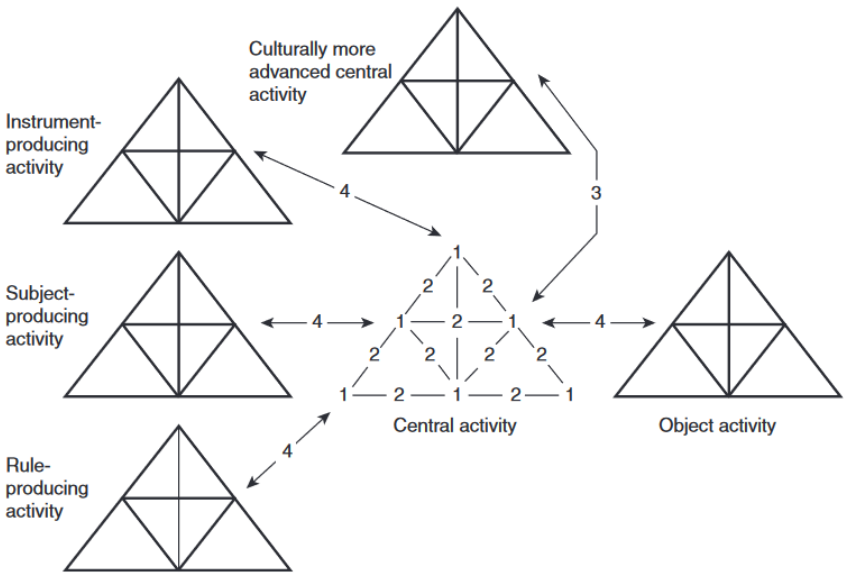


Figure 7.6 The four levels of contradictions (Engeström, 1987, p. 103).

The goal is to understand the future context and develop a conceptual representation of how the socio-technical activities discussed earlier in my research interplay in their network, leading to a better understanding of the challenges from multiple perspectives.

The following activity systems' sections gradually lead to the requirements used in the system sketches in Section 7.2. Sections 7.1.6 and 7.1.7 explore on a micro-level the activity systems' technological and social aspects, real-time and asynchronous collaboration, and their contradictions. Section 7.1.8 is based on Engeström's (2001, p. 145) case of researching two interacting activity systems. These two systems and their table of contradictions lead to a conceptual model for WikiSync. Lastly, Section 7.1.8, which is based on Blin & Appel's 2011 study of the classroom collaborative writing activity system's transformation, proposes a macro-level network of activities for the WikiSync's conceptual model. The network of activities is triggered by the PD design activity, leading to the training event using the WikiSync tool that employs real-time collaboration in alignment with the Wikipedia community.

7.1.6 Newcomers Real-time Editing Activity System

This section explores the Activity System of a hypothetical group of newcomers trying to edit Wikipedia in real-time after joining it with a training event. First, it reviews previous studies that analysed Wikipedia's editing Activity System to help guide the hypothetical activity illustration shown in Figure 7.7. This is followed by an examination of the contradictions between the components leading to implications, insights**,** or design requirements outlined in Table 7.1.

Studies, such as Slattery's (2009), have extensively analysed each component of editing Wikipedia articles, helping design activity systems for my research. For example, the author categorises the rules and division of labour component of Wikipedia's social infrastructure into official (Wikimedia's board of directors, policies) and unofficial (self-assigning a role to take care of a specific article overtime or answer other Wikipedia questions). In addition to Smith's study, other researchers, such Al-Ali (2020), guide the systems approach in documenting the elements of the activity. For instance, in a similar way to Al-Ali's (2020) approach to analysing a classroom activity's components, my research considers "time" as one of the non-physical artefacts. Since previous experience is also a mediator in shaping the Wikipedia article, my research merges the "experience" element with "time" for volunteering in Wikipedia as an artefact resulting in "time to volunteer by contributing an experience or knowledge" since they are connected and affect the object together.

AT is known to support researchers in understanding their complex research environment, asking informed questions, and communicating results, but it does not offer particular problems ready-made solutions that can be directly employed (Kaptelinin et al., 1999, p. 32). This has led many researchers to develop methods to bridge this gap between theory and practice by applying activity theory in domains such as HCI and eliciting IS requirements. For example, researchers such as Turner and Turner (2001) use an Activity Checklist analytical tool developed by Kaptelinin et al.'s 1999 combined with scenarios to provide a practical case of requirement elicitation using Activity Theory. They analyse the contradictions in an administration system that support student enrolment activity to argue that system contradictions can be opportunities and resolving can be utilised as the foundation for a "user-centred" new IS design. They use scenarios as a tool to reflect on data gathered from stakeholders and as a way to structure their information.

Unlike several previously mentioned studies, the description of the activity's sub-systems and their elements here will be limited to what is being discussed in the system's contradictions. The description will focus on the depicted contradictions, shown in Figure 7.7, that could occur internally (primary) and between the activity system components or nodes (secondary). Some elements will be discussed in accordance with their relevance to the study.

The activity systems used in my PhD research attempt to communicate the scenario through activity systems. They are based on the observational notes discussed in Chapter 4 and the sessions data analysed through thematic analysis in Chapter 6. In addition, the systems benefit from some of the concepts representing similar contradictions in previous work on

Wikipedia, using this framework (Bryant et al., 2005; Petrucco, 2010; Slattery, 2009; Hautasaari, 2013; F. Liu et al., 2020).

The first activity system shown in Figure 7.7 uses Engeström's first model to focus on an individual activity system as a unit of analysis. Similar to Turner and Turner's "what-if" scenarios approach is the activity system in the figure which is based on the following question: "what if a group of experts decide to volunteer their time editing a Wikipedia article akin to the way they collaborate in work with real-time collaborative writing tools". It maps how this group of newcomers cannot create a Wikipedia article using real-time collaboration without training due to tensions in the activity system, which is analysed later for solutions.

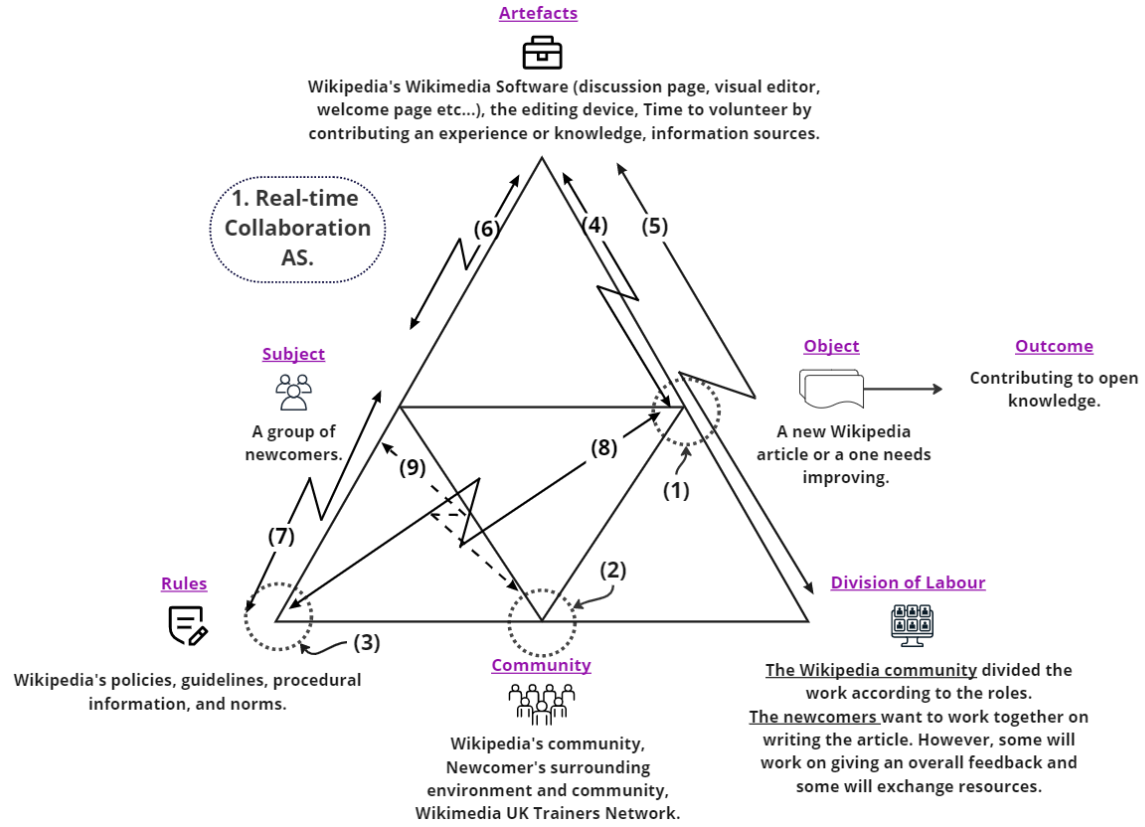


Figure 7.7 Current Newcomer's Real-time Collaborative Wikipedia Editing Activity System adapted from the structure of a human activity system (Engeström, 1987, p. 78).

The activity depicts a group of newcomers possessing different levels of expertise (subjects) who are motivated to learn how to contribute their knowledge (outcome) to Wikipedia through editing and submitting articles together in real-time (object).

Table 7.1 attempts to capture these contradictions and translate them into implications, insights, or design requirements to help avoid or ease some of the tension in the system interface design. This table is an adapted version of Turner and Turner's table "Contradictions,

Resolutions and Requirements" (p. 11.) In the "ID" column, the first digit refers to the activity system (AS) shown in Figure 7.7 and the second to the contradiction number in the figure.

Table 7.1 Contradictions in Wikipedia in real-time without training activity system.

ID	Contradiction	Implications, insights, or design requirements
1.1.	Primary Object: between the use value of the volunteering experience vs the exchange value of the time spent on volunteering.	Is going through this activity of learning and contributing to open knowledge worth the volunteer's time, which could be spent on another cause or the same cause of supporting open knowledge with a different system and bring me more gratification? **Insight.** This emphasises the importance of designing a system that facilitates volunteering opportunities that help the volunteers make the most of their volunteer experience, especially those with limited time or Internet access. As Shirky (2008, p. 253) puts it: "With the right kinds of collaborative tools and the right sort of bargain with users, it is possible to get a large group working on a project that is free for all."
1.2.	Primary Community: Contradicting culture of collaborative work and coordination vs culture of "lone wolf" and work in isolation.	In answering the questions on the use of real-time collaborative writing in my research, some respondents indicated that they would be uncomfortable working with others for such reasons as sharing the text in progress in real time could reveal some embarrassing mistakes and conflict with other editors working asynchronously on Wikipedia. **Requirements**. This has led to two design requirements, as will be seen later. First, a private editing space must be provided for collaborators. Second, the real-time collaborators' community work should not disrupt or affect the asynchronous one.
1.3.	Primary Rules: There are rules that anyone can edit, and Wikipedia's fifth pillar, "Wikipedia has no firm rules," vs extensive explicit and implicit rules that might complicate newcomers' lives.	Abiding by Wikipedia's rules is challenging for newcomers and even experienced editors (Reboot, 2017; Gluza et al., 2021). One of the most famous Wikipedia rules is that anyone can edit. But that contradicts the lengthy rules and complex community standards, guidelines and policies that make it hard for newcomers with different expertise to navigate the path to becoming Wikipedians (as discussed by Participant D in Section 5.4.1). This is apparent in some deletion policies that leave a newcomer with feedback that contains many community abbreviations and technical language. **Insight.** This emphasises the importance of designing a system that gradually equips newcomers with what they need for their journey.
1.4.	Secondary Tools – Object: The wiki tool contradicts the activity's object as the current tool	The technological ecosystem around Wikipedia is changing, and the relationship between the tools and the object has transformed over time (*Real-Time Collaboration Timeline - MediaWiki*, 2022). The accumulated tension between object and tool is rising. Over time, the Wiki tool no longer seems to correspond to

	doesn't allow producing articles edited in real-time.	the ever-evolving new opportunities in collaboration tools made possible in social media platforms and the workplace. **Insight.** This contradiction is one of the leading factors for proposing a more culturally advanced activity tailored to fit different groups' expectations in the collaborative writing stage, including real-time collaboration.
1.5.	Secondary Tools – Division of Labour: The division of labour to use the Wiki tool for real-time editing would result in edit conflicts.	Experienced editors have more influence over the norms, which over time, has become hard to change (TeBlunthuis et al., 2018). Adopting a new tool that could cater to the needs of this analysed activity's subject (the newcomers) is challenged by the current hierarchical division of labour. This tension between the Tools and DoL is similar to Bonneau's (2013) study of introducing new technologies for a university when the hierarchical DoL "*no longer appears to correspond to the new opportunities made possible in the workplace through the use of new tools*" (Bonneau, 2013, p. 3). **Requirement**. This contradiction is one of the leading factors in designing WikiSync as an external system to the Wikipedia community. The system could be tailored to allow more freedom in dividing the work, including co-authoring the same article simultaneously by actors who share the same power and influence on the environment.
1.6.	Secondary Tools – Subjects: Usability issues and lack of technical knowledge or writing skills to deal with the tools cause tension between the Subject and the Tools.	Through time, the usability of Wiki for editing degraded due to the difficulties in change discussed earlier. For example, even though the recently released changes to Wikipedia UI resulted from over a decade of work, hundreds of discussion pages about this change were still developing when this thesis was written (*Desktop Improvements - MediaWiki*, 2023). This has led to more attention in addressing some of the usability concerns that emerged in the sessions and more emphasis on designing self-explanatory interfaces for trainees with a wide range of technical skills.
1.7.	Secondary Subjects – Rules: The complexity of the rules frustrates some Wikipedia newcomers who have limited time. Some go through them and then edit Wikipedia. Others overlook them and start editing directly, leading to	This tension connects back to the multifaceted transparency theme from the thematic analysis (see 6.4.1). With time, the rules have helped organise the community of newcomers who started Wikipedia and governed its expanding boundaries of participants and content. The same community became what we now call the "core community" or "veterans". The historical development of the rules production activity became more centred around organising the work of this community and gradually became incompatible with and detached from supporting newcomers. Challenging, changing, or editing the rules to serve the analysed activity system's subjects in Figure 7.7 have become conflicting. Wikimedia Foundation and the community policies or guidelines

	community pushback.	such as the "Respect and Civility"[96] pillar, "WikiBullying",[97] "Be bold",[98] and "Please do not bite the newcomers",[99] among many others, tried to reverse the negative side of Wikipedia's editing strict standards and rules that are building up in complication over time. Such insight helps put into context some of the "Identify" session participants who have mooted the problems with onboarding. This is also a dilemma. Halfaker et al. (2011) note that reverts[100] (that remove someone's edits under specific rules) cause withdrawal in motivation to edit and communicate, primarily for "newbies" and experienced editors. Reverts also increase the edits' quality. On the other hand, Halfaker et al. find that editors who continue to do work in Wikipedia after having their work reverted enhance the quality of their future work. This is especially true for newbies and less productive editors. Halfaker et al. provide three design recommendations for social production communities 1) support communication, 2) encourage learning, and 3) focus on newcomers, which will be revisited in Chapter 8. **Requirement.** This tension has contributed to the design decision of having the platform as a mediator linking the Wikipedia community and newcomers. This system should create a safe environment that reinforces communication for the training activity subjects to collaborate, exchange personal feedback on work edits and learn gradually what is needed to contribute their expertise and be protected from some of Wikipedia's "*curmudgeonly old timers*" who "*should be kept away from newcomers until they have gained some experience in the system.*" (Halfaker et al., 2011, p. 172)
1.8.	Secondary Rules – Object: The current Wikipedia rules and guidelines disallow or discourage working on real-time articles.	Wikipedians on MediaWiki interpret collaboration as asynchronous (Collier & Bear, 2012). Therefore, their rules are designed for asynchronous articles, as the object of their collaboration activity. For example, editing in real-time is discouraged, as it would lead to edit conflict ('Edit Conflict - How to Guide', 2022). **Requirement.** This contradiction has brought on a design decision connected to contradiction 1.7.'s design decisions. It also led to the focus of the WikiSync design on creating new articles as the object of the activity, rather than editing ones that might be edited by the community members asynchronously. Of course, future iterations could look at different activities, such as editing existing articles, although this causes other tensions.

[96] https://en.wikipedia.org/wiki/Wikipedia:Civility
[97] https://en.wikipedia.org/wiki/Wikipedia:WikiBullying
[98] https://en.wikipedia.org/wiki/Wikipedia:Be_bold
[99] https://en.wikipedia.org/wiki/Wikipedia:Please_do_not_bite_the_newcomers
[100] https://en.wikipedia.org/wiki/Wikipedia:Reverting

1.9.	Secondary Subjects – Community: 1) Editing Wikipedia in a group setting creates tension for the Wikipedia community, as they grow suspicious of an intention to represent a biased or promotional view that motivates the activity. 2) Criticism of Wikipedia[101] causes some of the subjects' surrounding community to put pressure to question the value of the time spent volunteering on Wikipedia.	The first side of this contradiction connects back to the multifaceted transparency concept from Chapter 6's results. The second side is associated with one of the challenges newcomers faces and the "aha!" (or eureka) moment when they discover the amount of work the community put in to retain content reliability and that some of the criticism levelled at Wikipedia is unjustified. **Requirement.** This contradiction has engendered design decisions that allowed the community to view the articles in development to guarantee transparency and to provide an option of support if the subjects allowed it at that stage. The future design could look into benefiting from the change of newcomers' views by encouraging them to post their views on social media to counter the pressure put by their community, but this will not be covered in the first design for WikiSync, as it is beyond the scope of my research.

7.1.7 Two Interacting Activity Systems

The earlier activity covered a "what-if" scenario of using real-time collaboration in writing Wikipedia articles for a hypothetical group of newcomers without training. The following section, however, will utilise Engeström's (2001) new model for a "third generation of activity theory", as shown in Figure 7.8, which was introduced to bring an understanding of tensions in a network of interacting activity systems and to address "questions of diversity and dialogue between different traditions or perspectives." (Engeström, 2001, p. 135).

[101] https://en.wikipedia.org/wiki/Criticism_of_Wikipedia

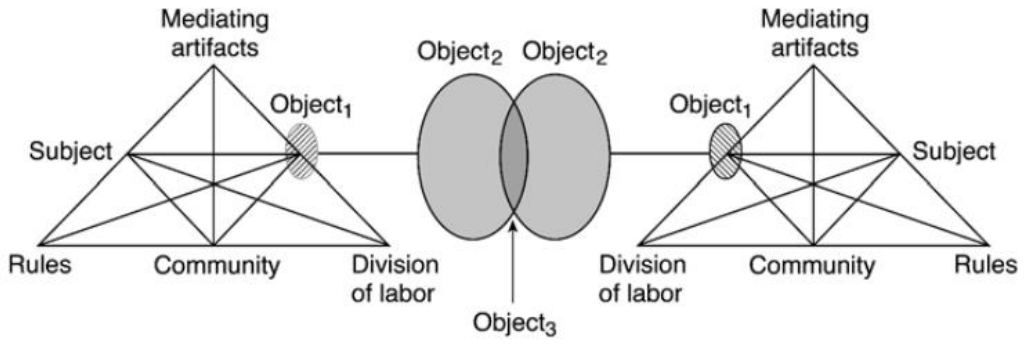


Figure 7.8 The third generation of activity theory for two interacting activity systems as a minimal model (Engeström, 2001, p. 136).

As shown in Figure 7.9 below, Engeström (2001) covers the case of public healthcare services in Helsinki, Finland, demonstrating this model's potential to find new work activity patterns to address tensions between two activity systems that share the same object.

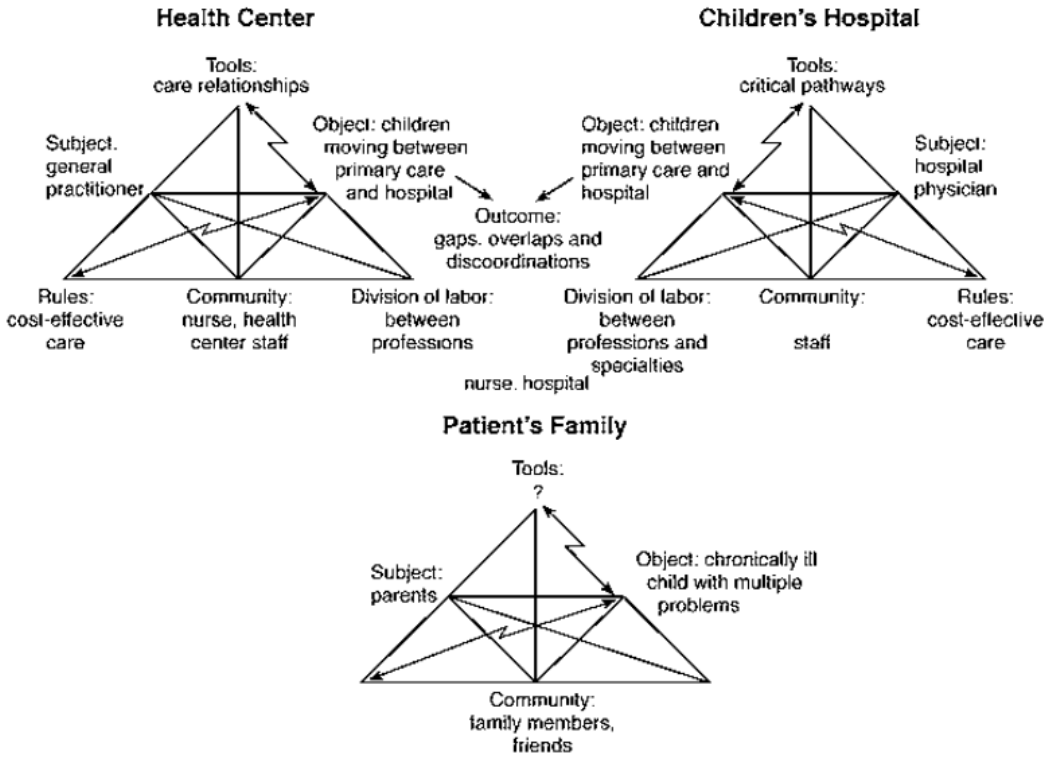


Figure 7.9 Contradictions in children's health care in the Helsinki area (Engeström, 2001, p. 145).

Other examples of this model include educational institutions using it to critically review the challenges and opportunities in the partnership and harmony between the university and schools in preparing teachers (Bloomfield & Nguyen, 2015). Bloomfield & Nguyen use the

activity theory as “a foundation lens” to link and connect the universities (who are preparing the teachers) and the schools (where teachers practice what they learned at university) for improved teacher integration and professional development.

The emerging findings from my research, combined with learnings from Engeström, have led to a focus on studying the following two interacting activity systems that share the same object (Wikipedia articles produced asynchronously):

- Activity System for trainers supporting newcomers with articles produced by asynchronous collaborative writing.
- Activity System for the Wikipedia community to review the newcomers’ articles to ensure they meet the community standards.

Figure 7.10 shows how these systems interact, pointing out the points of contradictions covered in greater detail in Table 7.3. Even though the two systems share the object of collaboratively constructed knowledge that would be shared to benefit the world, many contradictions in the systems produce an unexpected outcome of a high drop-off rate and lack of continuous collaboration.

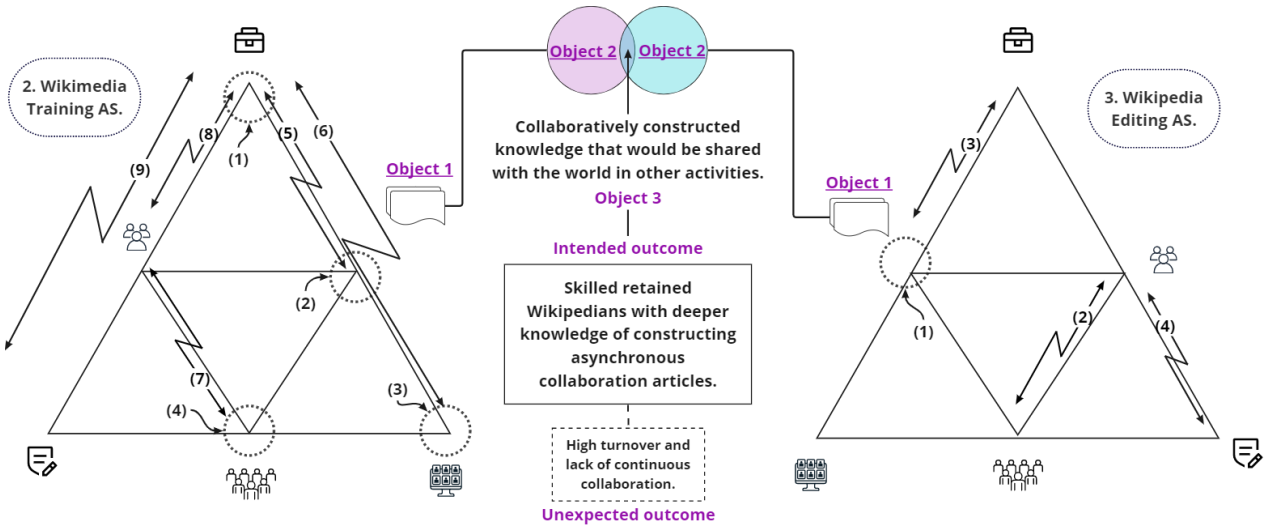


Figure 7.10 Two interacting activity systems for trainers Wikipedia community adapted from the two interacting activity systems as a minimal model for the third generation of activity theory (Engeström, 2001, p. 136).

Identifying contradictions in these two interacting systems contribute to uncovering the implicit challenges for newcomers and introducing real-time collaboration in training environment activities using WikiSync. Table 7.2 compares the elements of these two activity systems that are shown in Figure 7.10 based on Engeström’s (2001) model. The following

Table 7.3 covers the numbered contradictions. Section 7.1.8 elaborates on these two systems' interrelations as part of a network of activity systems.

Table 7.2 The elements of two interacting activity systems shown in Figure 7.10 based on Engeström's (2001) model.

AS Element	Wikimedia Training AS	Wikipedia Editing AS
Subjects	Trainer, co-trainers, trainees (or newcomers).	Wikipedia community members involved in supporting newcomers and reviewing their work.
Mediating Tools	Training material, presentation tool (Google Slides),[102] Google Sheets for worklist with the work lists to assign the names of those who would work on each article and some links to web pages about editing guidelines, conference meeting tool (screen-sharing feature, Direct Messages (DM)), PDF document or link-sharing using DM, or Email, Wiki editor (example page - user page - sandbox), Programs & Events Dashboard.[103] Feedback, trainees' background. Time to volunteer by contributing an experience or knowledge.	Wiki software, discussion pages, watchlist,[104] Teahouse,[105] Libera Chat[106] network chat rooms, Wikipedia's resources and guidelines, administrators' tools (e.g., Twinkle,[107] Bots),[108] and Feedback.
Object 1	Trainees are onboarded with real-time coordination and ready for asynchronous collaboration through the trainer's real-time feedback or editing articles directly in their draft spaces.	Trainees are supported with asynchronous feedback on their articles.
Object 2	Refined articles accepted by Wikipedia's community.	Articles that meet the community standards.
Object 3	Collaboratively constructed knowledge that would be shared with the world in other activities.	
Rules	Training timing, training cultural norms and rules (e.g., Safe Space Policy),[109] assigning names on the Google Sheet's	A list of Wikipedia's community policies, guidelines and

[102] https://en.wikipedia.org/wiki/Google_Slides
[103] https://meta.wikimedia.org/wiki/Programs_%26_Events_Dashboard
[104] https://en.wikipedia.org/wiki/Help:Watchlist
[105] https://en.wikipedia.org/wiki/Wikipedia:Teahouse/About
[106] https://en.wikipedia.org/wiki/Libera_Chat
[107] https://en.wikipedia.org/wiki/Wikipedia:Twinkle
[108] https://en.wikipedia.org/wiki/Wikipedia:Bots
[109] https://wikimedia.org.uk/wiki/Safe_Space_Policy

	worklist, and Wikipedia rules (community's and Wikimedia's). The accepted number of trainers and co-trainers (or experienced editors) compared to the number of trainees, norms of welcoming and rewarding other contributors, and the trainees' cultural norms.	procedural information,[110] norms of welcoming and rewarding other contributors, office actions or Wikimedia's advanced rights.[111]
Community	Co-trainees, trainers, expert Wikipedians, trainees' and trainers' surrounding environment and community, and Wikimedia UK network of trainers.	Wikipedians, Village Pump community,[112] Wikipedia readers, Wikipedians' surrounding environment and community, trainers, Wikimedia, and public relations companies.
Division of labour	Trainees learn to edit wiki pages, interact with co-trainees in chat, trainer facilitates (feedback cycles on drafts and solve conflicts), Co-trainer support the trainer and trainees (Wikitext, rules, or technological issues), expert Wikimedian support and guide through providing feedback (sometimes experts join for support, as discussed in Chapter 4). Horizontal division of labour with each person working on a specific article or section of an article. Sometimes exchanging resources about the articles in hand (all subjects).	Wikimedia Foundation oversees the technological infrastructure, elected administrators oversee the community; volunteers with hierarchical permissions mostly contribute their area of expertise, while bots have the specific tasks to fix style or typos, and admins and experienced editors sometimes support trainers and collaborate with them in training sessions.

Many discussions have led to conceptualising the system as an external tool to reduce some of the challenges mentioned in Table 7.1. However, integrating the trainees into the Wikipedia community or their real-time-created articles using such an external tool will also result in the inevitable frictions discussed in previous sessions' insights. Table 7.3 will discuss more contradictions at the two activities level, leading to the new tool's system specifications that use the training environment to introduce real-time collaborative writing.

The discussion of the contradictions listed below is based on the ethnographic study, as well as the sessions analysis. Similar to Table 7.2, the first number of the "ID" column refers to the number of the activity shown in Figure 7.10, and the other one is for the contradictions.

[110] https://en.wikipedia.org/wiki/Wikipedia:List_of_policies_and_guidelines
[111] https://meta.wikimedia.org/wiki/Office_actions#Use_of_advanced_rights_by_Foundation_staff
[112] https://en.wikipedia.org/wiki/Wikipedia:Village_pump

Table 7.3 The contradictions between two interactive activity systems for a Wikipedia training session and the Wikipedia community support.

ID	Contradiction	Implications, insights, or design requirements
	2. Wikimedia Training Activity System (AS).	
2.1.	**Primary Tool**: Limiting tools used in training to those provided by Wikipedia vs using tools like Google for collaboration Doc for drafting and then moving to Wikipedia.	Limiting the trainees, as much as possible, to what tools Wikipedia offers could help get newcomers ready to be fully-fledged community members who are independent from tools the Wikipedia community does not use or even discourage using. On the other hand, training newcomers in a few hours could considerably benefit from a few minutes saved through tools that foster coordination or collaboration among participants. **Insight.** Learning how to edit Wikipedia is important, but some trainees might prefer getting help with Wikipedia through tools they are familiar with. This opens up the possibility of distributing the work among participants. Some will edit in WikiSync, while others interested in developing new skills would work on Wikipedia to "Wikify" the articles there.
2.2.	**Primary Object**: The newcomers' motivation to learn a new skill vs the motivation to support open knowledge.	This is a healthy tension, although the increased difficulty in learning the skill and using the tools (Reboot, 2017) has created secondary tensions, as well (see 2.5 and 2.8). **Insight.** This emphasised the same point in Contradiction 1.1 on the importance of designing tools that bring back balance and ensure that designing WikiSync would improve the volunteering experience.
2.3.	**Primary Division of Labour**: Working individually in an asynchronous way vs collaborating in real-time with other trainees.	As discussed earlier in Section 4.4.2, observation T1.O5, trainees are being asked to work separately to avoid edit conflict. Therefore, **the new system should** allow flexibility in dividing the work.
2.4.	**Primary Community**: Community as supporters for newcomers vs	The community has many challenges, as discussed in Chapter 2. The reduced volunteer number puts a lot of pressure on remaining community members, who face increased vandalism or attempts to penetrate their community by public relations companies to push particular political or commercial agendas. As a result, many rules and

	community as gatekeepers against vandalism.	implicit norms were introduced (Halfaker et al., 2013), contributing to a low-trust environment and suspicion of newcomers. The Wikipedia community gradually have less time or motivation to support newcomers' work at training events (Reboot, 2017; Gluza et al., 2021). This shows the importance of designing features that make the most out of having trainees work and revising each other's work before being checked by a trainer. Furthermore, **the system should** allow a straightforward process to deal with incoming work from the tool to reduce the tension in this node.
2.5.	**Secondary Tools – Object:** Many trainees are accustomed to editing with a Phone, which is not supported or covered by the current training tools.	During my observational study, several trainees could not finish the training, as they had joined from their mobile devices, and Wikimedia's training sessions were designed for desktops. **Insight.** Since my research is focused on providing a tool with real-time training sessions, this contradiction is out of the research focus. But as will be discussed in Section 7.3, the last phase shaped the design to be mobile-friendly.
2.6.	**Secondary Tools – Division of Labour:** The trainer, as a "reviewer" role, is unable to provide support on an article unless the trainee shares her/his screen, something that is not straightforward to all trainees with the current technology.	While observing the training's asynchronous collaborative editing, one participant edited her/his sandbox[113] space on Wikipedia and needed some help with the article. The trainer asked the trainee to share her/his screen to help them, which was technically difficult for the trainee. **The new system should** allow a trainer to access the developing article easily.
2.7.	**Secondary Community – Subject:** Lack	Trainers who do not declare their training of participants increase the chance of critical feedback from the Wikipedia community to the

[113] https://en.wikipedia.org/wiki/Template:Spaces/sandbox

	of coordination between the trainers and administrators. Systematic bias.	participants, or even blocking the IP of the training venue, levelling up the trainers' duties. Also, the trainees are advised to put a message on their user pages that include text along the lines of "Please do not bite (link to the article mentioned earlier)[114] me. I'm a newbie! My name is [Your Name], and I am a [About You]. I am taking part in training to learn to be a Wikipedia editor. This page is a work in progress, and I am approaching my subject in good faith (linked to the Assume good faith article).[115] Thanks!" In turn, the caution in the recommended message could make some participants tense about what will come after. Systemic bias by administrators toward content from subjects could lead to a breakdown in this activity. **Insight.** This connects to some of the earlier design decisions in ensuring transparency (see 6.7.1) that could build trust for trainees coming from a WikiSync system.
2.8.	**Secondary Tools – Subject:** 1) Trainees lack technical knowledge or writing skills to deal with the tools, 2) trainees and trainers have the challenge of finding resources that fit Wikipedia's rules of notability,[116] 3) trainers cannot check the chat and demonstrate at the same time, 4) trainers need for a new account to demonstrate to trainees.	Many challenges and tensions exist between these two components of the training activity system. For example, the interfaces for Wikipedia change once someone creates an account and sets the preferences. Once created, the trainer's user page link would turn blue instead of red. The trainer's account, in this case, cannot be used to demonstrate to trainees how to go through certain steps. Also, Wikimedia's tactics of rolling out changes start with new users' accounts only as a way of testing and not changing the old user experiences, making it even more difficult for the trainers to use their accounts for demonstration. If not done correctly, creating another account for demonstration could lead to blocking both of the trainers' accounts for Sockpuppetry[117] policy violation. Not all those who train know they can create another account for such training purposes. **The new system should** provide the trainers with flexible demonstration features and experimental Wiki Space that simulate Wikipedia's interfaces.

[114] https://en.wikipedia.org/wiki/Wikipedia:Please_do_not_bite_the_newcomers
[115] https://en.wikipedia.org/wiki/Wikipedia:Assume_good_faith
[116] https://en.wikipedia.org/wiki/Wikipedia:Notability
[117] https://en.wikipedia.org/wiki/Wikipedia:Sockpuppetry

2.9.	**Secondary Tools – Rules:** Tools that support the current and experienced Wikipedians could be used against minority views, which is contrary to some of the openness community norms and rules.	The design of Wikipedia tools and processes have "built-in (value) biases [There are]tensions between the stated values of Wikipedia [...] and the values in practice within the relatively small and homogenous group of core Wikipedia contributors, who sometimes prioritise values, such as freedom from censorship over multicultural inclusivity." (Morgan et al., 2012, pp. 3490, 3495). The authors also note that sometimes the rules are being invoked to impose an opinion rather than inform the contributors. Kriplean et al. (2007) refer to this as "power plays" caused by ambiguity in policies that need to be taken into account when designing tools that support mass participation. This connects to what is discussed in Chapter 4 with some participants who feel that the administrators' feedback is critical, which, in turn, shares similarities with Gluza et al.'s (2021) observational study of Edit-a-Thon findings. They discuss how the discussion of participants in a training event with an administration over notability rules drove them "frustrated and culminated to a point where they themselves were in doubt whether they could bring any change to Wikipedia and whether it would not be more effective to leave possible changes to Wikipedia for the administrator to do." (Gluza et al., 2021, p. 7). **Insight.** Understanding the power dynamics that affect the usage of tools helps in designing tools that alleviate the problem rather than contribute to it. As will be discussed at the end of this section, studying this tension leads to more caution in devising the conceptual model of the new tool, which contributes to reducing the tension, instead of putting the WikiSync trainees in confrontation with the community.
3. Wikipedia Editing AS.		
3.1.	**Primary Object:** Trainees supported by the Wikipedia community with asynchronous personalised feedback exchange, creating a welcoming environment for newcomers to edit articles vs newcomers that are provided with concise, timely	This tension shares similarities with the discussion over contradiction 2.4, which questions the community's level of engagement in supporting the training articles outcome. **The new system should** enable the community to give early feedback that may save time and prepare trainees ready to be integrated. However, this should be approached cautiously, as early feedback could be destructive rather than constructive. Gluza et al.'s (2021, p. 6) note that "throughout the Edit-a-Thon the administrator enforced rules regarding verifiability and citations, copyright of material or notability of a person very strictly which caused frustration among editors who already struggled with conceptual challenges of that kind when contributing to Wikipedia." **The new system design should** allow feedback when the trainer and trainees are ready. It should have a feature allowing/recommending administrators and experienced editors to provide constructive feedback. This will not be covered in depth in WikiSync design, as it is beyond the research aim.

	feedback to protect Wikipedia quality and invest the time in other areas.	
3.2.	**Secondary Subjects – Community**: As time passes, the core Wikipedia community members are facing more challenges which led to stricter behaviour and other tensions mentioned earlier.	Assumptions about Wikipedia and its editors, criticism from the readers and academia regarding the articles' quality and increased vandalism and attempts of paid editing by PR companies to push their agenda on some articles are some of the few contributors to stricter behaviour and other tensions mentioned earlier. The larger the challenges, the more the community is susceptible to losing volunteers, and the more rigorous the regulations and reliance on bots to reduce pressure on remaining volunteers lead to frustration by newcomers trying to join the community. This creates a vicious cycle related to tensions in the rule's component. **Insight.** This emphasises the importance of features that build trust and transparency between the WikiSync community and the Wikipedian community for a constructive complementary relationship with onboarding newcomers.
3.3.	**Secondary Tools – Object:** Relying on the bots that increase the chance of false positive that results in article deletion with automated messages.	This connects to contradiction 3.2, as discussed earlier. **The new system should** have a combination of case-by-case feedback, automated tips and notes that flag particular concerns when editing to avoid getting the articles filtered out through the community bots.
3.4.	**Secondary Rules – Subjects:** Spending too much time on external tools or talk pages to guide newcomers will not give the community	Wikipedia community members are proud of their edits, or the articles they have created that contribute to getting the "trust and confidence" needed to apply for higher ranks, such as "Administrator" (*Wikipedia Administrators Policy*, 2022). However, for some, spending time on talk pages and helping in the training session will not count toward their fellows the same way contributing to articles do. This creates a challenge in getting the support of administrators. **Insight.** In a new system, such as WikiSync, that would be solved had a way been found to reward the experienced who help in WikiSync on

	members the same status as contributing to articles.	a dedicated Wikipedia page designed for WikiSync to show and celebrate the impact they are making.

This is an attempt to form a view of the complexity of change in Wikipedia when the history is considered of how its rich socio-cultural and socio-technical infrastructure is shaped through time to make better design decisions for its future.

Having had more time, looking at the activities that produce some of these rules or tools in their context and activity system would be interesting. For example, in addition to the listed contradictions, it is important to see how the community's feedback as a tool in Wikipedia Editing AS3 is being produced in its own activity system (subjects, tools, object, community, division of labour, rules). The trainees in Wikimedia Training AS2 are provided with a safe space for newcomers. It encourages them to write their articles, which receive detailed, constructive feedback in real-time. In AS2, the trainer's feedback is considered a tool for a successful activity. On the other hand, this contradicts AS3's usage of feedback as a tool where the feedback could be straightforward, direct or rough, and uses vocabulary common to experienced Wikipedians and motivated by high standards, which may be accompanied by a Speedy Deletion[118] action.

Back to the research focus on real-time collaboration, trainees' experiences with the rules in activities AS2 and AS3 are contradictory as the use of real-time tools, such as Google Docs, could exist in AS2 (as discussed in Chapter 4), yet it does not exist or contradict with the AS3 activity components. In other words, unless hacked around the system, those who use real-time collaborative editing with their trainers in AS2 find themselves in AS3, forced to adapt to the "liquid collaboration" model.

Conceptual Model of Real-time Collaboration Using WikiSync

Many systems and tools (social and technical) are at play in the discussed Activity Systems. The observation study and sessions' analyses using the AT help highlight real-time collaborative editing challenges and give insights into how to design WikiSync.

Examining how AS2 and AS3 interplay, it is evident that the new real-time collaboration activity should be viewed as part of a network of systems and considerations. However, before building a network of activities, I had to clearly define how real-time collaboration could be

[118] https://en.wikipedia.org/wiki/Wikipedia:Criteria_for_speedy_deletion

introduced using systems that lead to the design of the interfaces. Therefore, a literature review exists to identify how these different concepts and systems could be linked to form a conceptual model that acts as the foundation for designing WikiSync. First, Zurita & Nussbaum's (2007) model of Computer Supported Collaborative Learning (CSCL) systems, shown in Figure 7.11, has helped translate the ideas about activity components, insights, and requirements into their position as software. The model is designed for a group collaborative learning experience using a mobile device to help distinguish the characteristics of each component. For example, looking at the tools and community crossing boundaries and interacting with each other to form the network has led to the idea of forming a community interface, which is discussed later in this section.

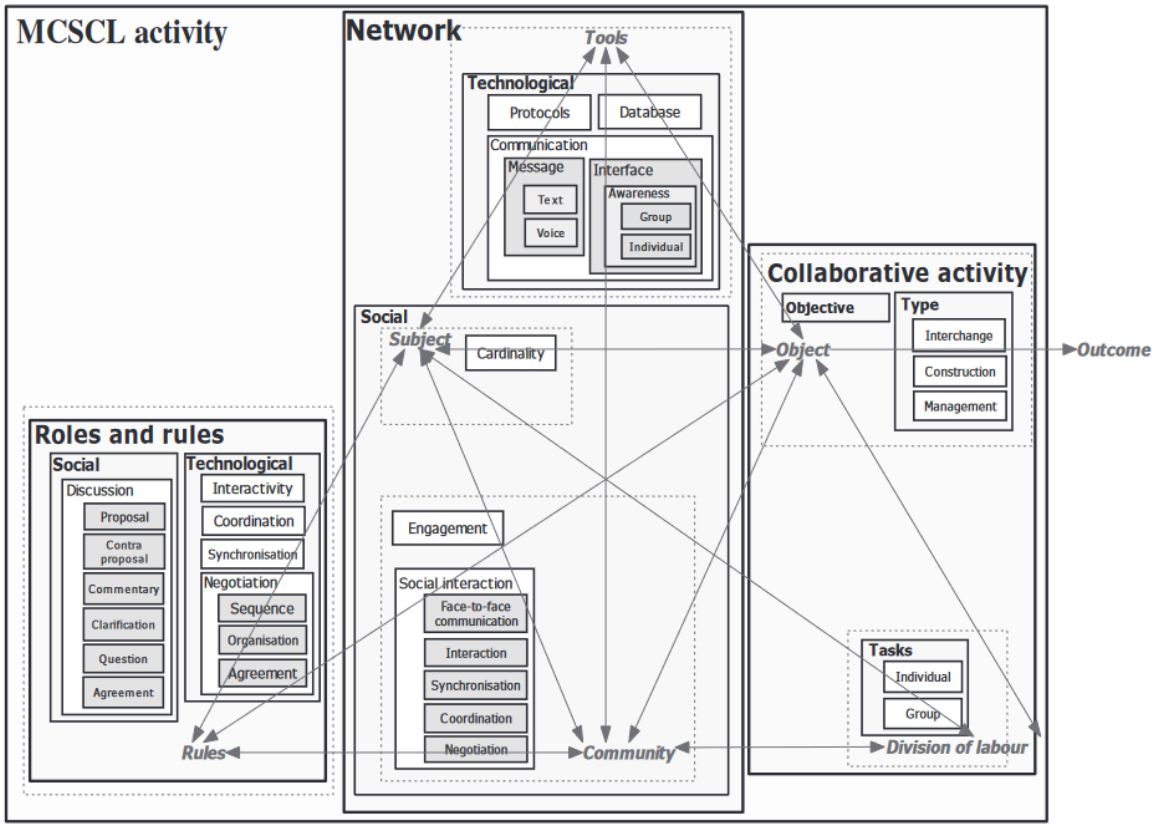


Figure 7.11 Mobile computer-supported cooperative learning (MCSCL). Based on Engeström's enlarged activity theory model (Zurita & Nussbaum, 2007, p. 217).

At this stage of the research, there is an emergent need to revisit the concepts discussed about the real-time collaboration feature in Wikipedia training and see how this new understanding of the tensions and interplay between the various parts of the WikiSync tool would translate into a conceptual model.

The aim is to position WikiSync as a mediator linking the Wikipedia community with newcomers, providing a transparent, innovative, smooth transition from the way WikiSync is used to the methods of work in the Wikipedia community, which can use this tool proactively to prepare the next generation of highly-skilled, retained Wikipedians that possess more profound knowledge on how to collaborate on constructing knowledge that would benefit and have a greater impact on the world.

The conceptual model addresses critical social challenges that have emerged from the discussion. The most critical concerns, discussed in detail later, are:

- The need for real-time collaboration transparency;
- The high possibility of getting the article content written in real-time rejected by the Wikipedia community;
- Or even getting the trainees' IPs blocked.

Thus, co-designing an external tool that functions as a middle layer between Wikipedia and the training participants (newcomers and trainers) is discussed. The tool's concept is divided into master interfaces, systems and actors, as shown in Figure 7.12 and discussed below.

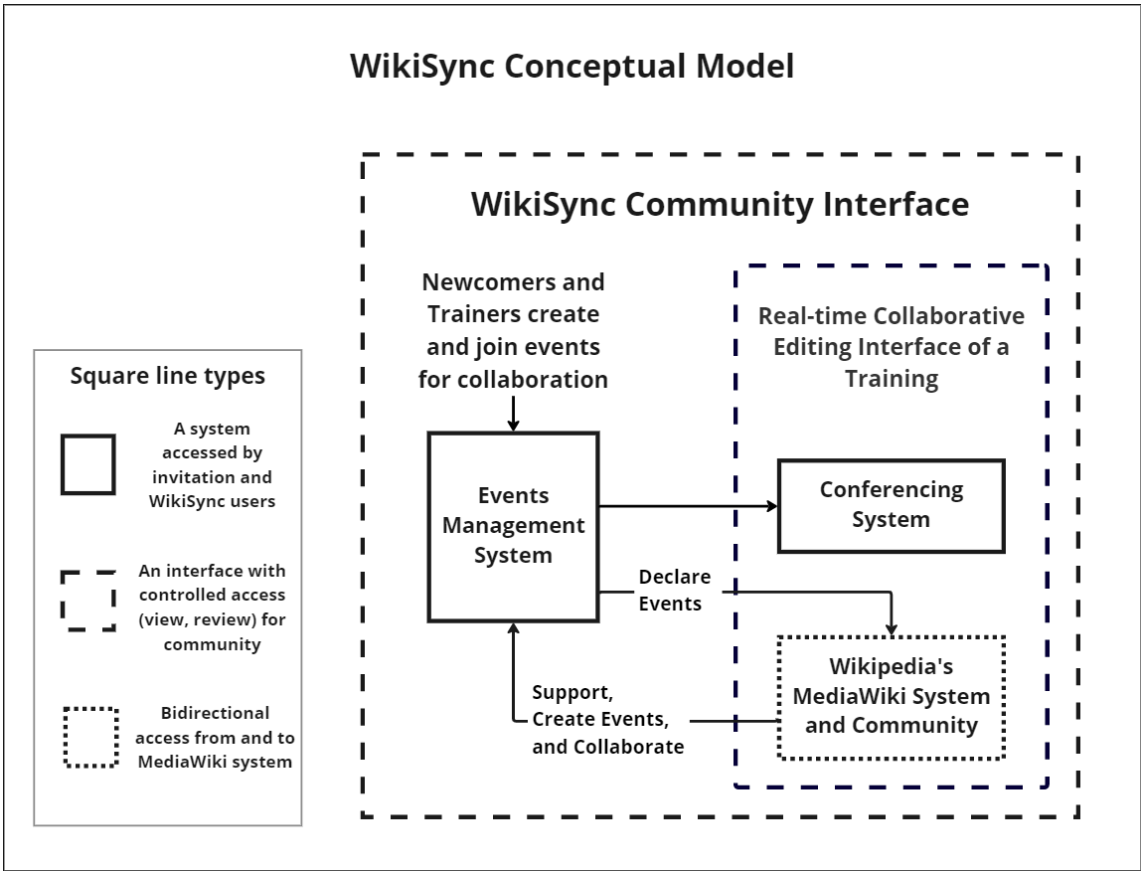


Figure 7.12 WikiSync conceptual model.

As Wikipedia training involves either one or two parts -a general introduction to Wikipedia and a writing part, a training event on WikiSync could be a training session that might include writing articles or a writing event without training. The focus is on an event that has both parts.

After an event is declared on the WikiSync community interface, it should be listed on a Wiki page (possibly under a Project page on Wikipedia). Wikipedia community members could help by participating as co-trainers through the event or as participants in the editing activity, where they can edit and review articles that end up on Wikipedia.

The WikiSync community space has an event management system enabling trainers to design, edit, publish, and run the training event. It is where community members can review events, express interest, or join in response to an invitation.

This management system generates training events facilitated by a conference call system. If the training includes editing, then attaining this activity would lead participants to access a real-time collaborative editing interface for the purpose of collaborating on editing articles. Each article being worked on would be supported by its dedicated conferencing call and asynchronous chat.

Furthermore, WikiSync allows for transparency and flexibility in getting Wikipedia Community members to access (when allowed by trainer) and support the training. It would be possible to send the content produced in this real-time collaborative interface back to be reviewed by the WikiSync community, as well as Wikipedia's community on the MediaWiki system. Also, on request from the training participants, community members can directly review a developing article section by clicking on the notification listed in the "latest activities" section of the WikiSync community interface.

The process involving the WikiSync event, which is discussed in Section 7.2, highlights how the Wikipedia community can play a significant role in the WikiSync community by supporting its trainees.

It is worth noting that such a solution that calls for introducing the change through an external tool does not necessarily align with Wikipedia's approach of a unified platform that caters to the needs of all communities in one place, which in most cases promotes standardisation and centralisation leading to more consistency in the user experience. However, such an approach could overlook specific needs or cultural nuances. In this thesis, I argue that creating multiple external, specialised systems that support or feed into Wikipedia can better cater to its community's diverse needs for more customisation than the unified platform approach. There are already several similar systems, such as Wiki Education Dashboard, which is a project by Wiki Education, a small non-profit spin-off of the WMF. Since 2010, this

project, which is external to the Wikipedia platform, has represented an external influence on the platform by training thousands of individuals who added over 100 million words to Wikipedia (*Wiki Education Impact*, 2022). However, currently, most of these are the result of one-off success stories rather than a user-friendly systemic mechanism to create an ecosystem of experimentation and innovation surrounding Wikipedia that could transform into a change in this unified platform.

Nevertheless, the fragmentation of the system design approach could lead to inconsistencies and duplication of efforts. Therefore, such an approach requires high levels of coordination and collaboration between the system stakeholders, leading to higher ownership and easier change processes. This thesis provides a clear example of such an approach.

7.1.8 WikiSync Activity Systems Network

The last step is to design the new activity system of real-time collaborative editing as part of training. As discussed in principles 1 and 5 in Section 7.1.4, the activity should be considered within a network of activity systems that develop as a cycle that starts with individuals questioning contradictions, which in turn, set in motion changes to other activities in its network, encompassing a wide range of possibilities for change and improvements. Therefore, it is not only about real-time collaboration activity, rather, the focus should also be on how this activity functions among others.

The work in WikiSync lies in between two types of computer-supported systems that support collaborative learning and writing. The previous section already covers Zurita & Nussbaum's (2007) work on CSLS and helps build the conceptual model of WikiSync. This section builds on that by exploring Blin & Appel's (2011) study, in which Activity Theory is used to analyse collaborative writing activities in second language teaching. Blin & Appel's model (shown in Figure 7.13 depicts the transformation of the Activity System in collaborative group writing.

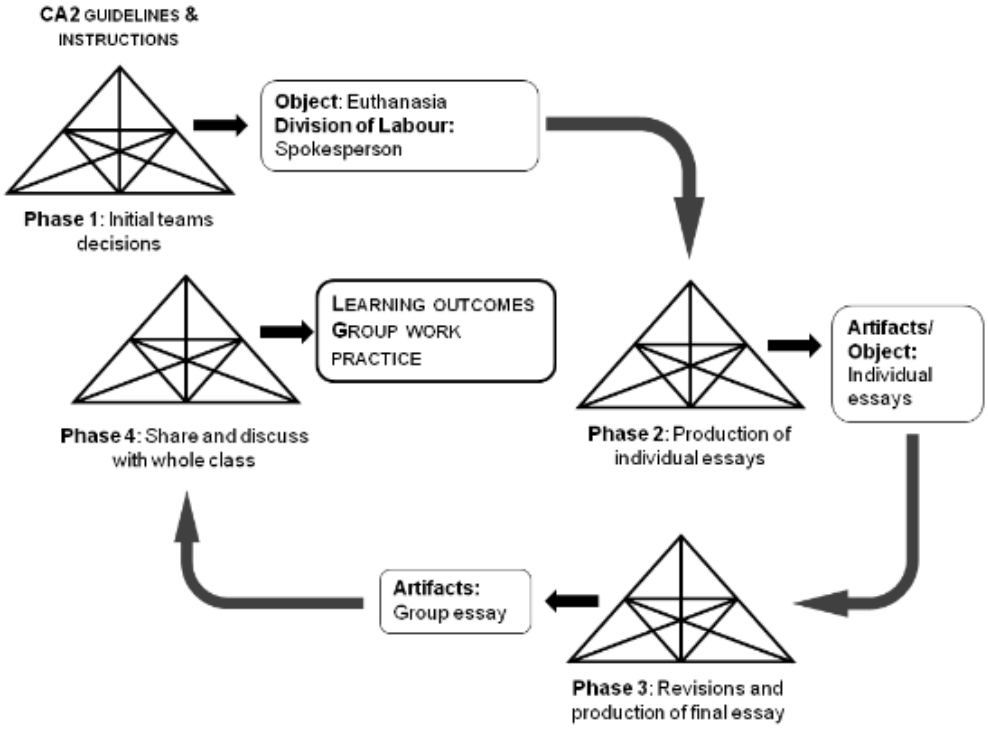


Figure 7.13 The transformation of an online collaborative writing activity system for student group assignment (Blin & Appel, 2011, p. 484).

Introducing real-time collaborative technology should be framed in a socio-technical model that would allow such technology to be accepted by the community. Therefore, my research had to critically examine the training to identify the challenges and opportunities to introduce said technology.

The earlier understanding of the multilevel of contradictions, combined with Blin & Appel's model, has engendered the design of the conceptual network of activities shown in Figure 7.14. All the activities in the diagram are considered as separate activities with their own socially recognised separate objects of activity rather than actions whose objects may serve multiple activities. Since this work focuses on real-time collaboration, the figure's activity systems AS2, AS6, and AS7, coloured in red, are identified as the three most critical to the WikiSync tool. The following Activity Systems' components are covered in detail in Table 7.4:

1) AS2 -Training activity system with trainees ready to write articles in real-time (RT).
2) AS6 - RT collaborative writing activity system with real-time created articles as an object.
3) AS7- RT revising activity system with articles ready for publication as the system object.

The introduced conceptual network of activities in Figure 7.14 first depicts my research's PD process as the initiating activity on the left, which provides a new instrument and a usage scenario to AS2 in red, on the left. The second real-time collaboration training activity prepares the subjects for a real-time collaborative writing activity. AS3 is when the training participants agree on the real-time collaborative writing plan. This activity's object transforms into an instrument in AS4, where participants review and research the articles' list of sources to write from. Moreover, it may involve subjects from the Wikipedia community who opt to help with the training. The object from activity AS4 becomes the instrument in writing activities AS5 and AS6, whose objects are articles ready for revision by subjects in other activities. AS7 involves the Wikipedia community, as well as trainers and trainees, in revising the articles produced, which are transformed into a collaboration activity for the trainer with the community to publish these articles to Wikipedia. Finally, the published articles become part of a Wikipedia community's collaboration process to ensure that they meet the guidelines and adhere to community rules, and by extension, guide the newcomers through messages, notes, and feedback left on their newly-created articles or user talk pages.

This network of activity systems, starting with the activity produced by the PD design process, would lead to a tertiary contradiction between the training activities and the old training process. However, this is predictable and has been covered in the earlier sessions when we discussed the challenges of real-time collaboration.

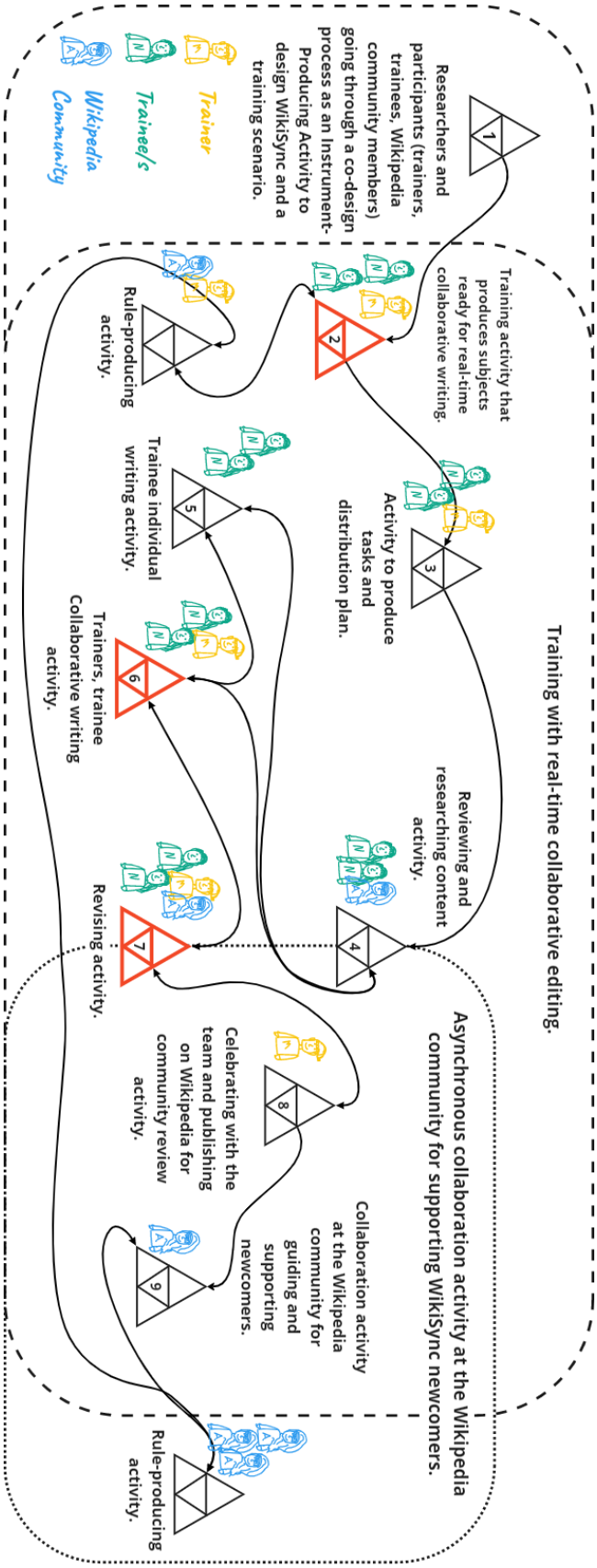


Figure 7.14 Conceptual network of activities depicts the sub-activities of a training activity system that uses real-time collaboration triggered by the PD design activity and adapted from (Blin & Appel, 2011, p. 484).

Table 7.4 The elements of three activity systems highlighted in Figure 7.14 and based on Engeström's (2001) model.

AS Element	AS2 - RT Collaborative Training	AS6 – RT Collaborative Writing AS	AS7 - RT Collaborative Revising AS
Subjects	Trainer, trainees (or newcomers).	Trainer, a small group of trainees.	Trainer, a small group of trainees, Co-trainer, Wikipedia's community.
Object	Trainees know how to collaborate in real-time on writing articles.	A new Wikipedia article written through real-time collaboration.	Review new Wikipedia articles that meet the community standards and articles ready for publishing
Mediating Tools	WikiSync editing space. Conferencing system with document sharing and DM features. Trainer's training material and planned activities. Presentation tool. Editing device.	WikiSync editing space. Conferencing system with document sharing and DM features. Feedback Time to volunteer by contributing experience or knowledge. Collected resources for writing. (Books, tips, videos, newspapers)	WikiSync editing space. Conference meeting tool with document sharing and DM features. Feedback Collected resources for writing. Trainees background.
Rules	Training event management (starting the training, timing, and access control). Training cultural norms. Wikimedia Foundation rules (Safe Space Policy) Assigning group names on the collaborative	Wikipedia guidelines and rules shared in the training. (Collaborative/Private spaces, how many people can work on a section, who can add new sections, who can submit an article as ready for review or for	Wikipedia guidelines and rules shared in the training (rules for allowing the Wikipedia community restricted access to revise). Team initiation and task distribution plan activity.

	articles for task distribution. Wikipedia guidelines. The number of trainers compared to the number of participants. (From the observation study, if the number of trainees exceeds 12, the trainer needs to find a co-trainer to help) Norms of welcoming and rewarding other contributors. The newcomer's cultural norms.	publishing on Wikipedia) Team initiation and task distribution plan activity.	
Community	All training participants (trainees, co-trainers, expert Wikipedians) Trainees' surrounding environment and community. Wikimedia trainers' network. Wikimedia Foundation.	All training participants (trained participants, co-trainers) Trainees' surrounding environment and community. WikiSync community.	All training participants (trained participants, co-trainers) Trainees' surrounding environment and community. WikiSync community.
Division of Labour	Trainers train the participants. Co-trainer support (text, rules or, technological issues).	Trainees write and edit articles collaboratively. Co-trainees watch and interact in the chat. Trainer facilitates (Feedback cycles).	Trainees revise and edit articles collaboratively. Trainers, co-trainers, and newly trained Wikipedians in WikiSync comment and edit other trainees' written sections.

			Co-trainees interact in the chat. Trainer facilitates (Provide feedback cycles on drafts and solve conflicts) Co-trainer support (text, rules or technological issues) Wikimedian support and guide through feedback and sometimes Exchanging resources (Wikipedians).

7.2 Low Fidelity Sketching

Using AT as an analytical lens has helped identify the WikiSync conceptual model and how the different activities can interplay to introduce real-time collaboration. The plan was to build 1) a storyboard, 2) a business process of the WikiSync system, and 3) low-fidelity sketches that could be shared with the participants in the "Create" session of the Design phase to evaluate and iterate on the solution we have collaborated on so far.

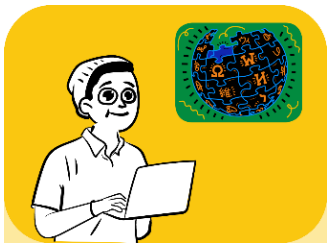
System Usage Scenario

Following the discussion in Section 7.1.6 about adhering to other research approaches, such as Turner & Turner's (2001) in developing usage and "what-if" scenarios, and Matthews et al.'s (2011) "collaboration persona", my research has benefited from this technique in communicating the AT analysis findings to my participants in a "what-if" scenario for a collaborative persona.

A scenario has been designed here that starts with a hypothetical team of librarian colleagues. The group characteristics are discussed earlier (see 6.2.3).

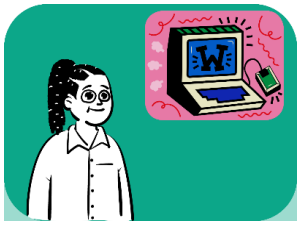
The Full Scenario of The First Training Event for WikiSync Participants. [119]

"WikiSync participants" refers to all who participate in the training event, i.e., trainees, trainers, co-trainers and Wikimedia community members offering support.



#1 Mobo is a 33-year-old Nigerian university lecturer who speaks English and Yoruba and has a bachelor's degree in computer science. He is interested in being part of the global and respected community of Wikipedia to collaborate on creating reliable knowledge that would help him learn about new cultures, as well as help others. Teaching editing Wikipedia is Mobo's way to encourage forming neutral, reliable knowledge for the world. After five years of editing Wikipedia, he received online Training from the Wikimedia UK Charity two years ago and only delivered online training.

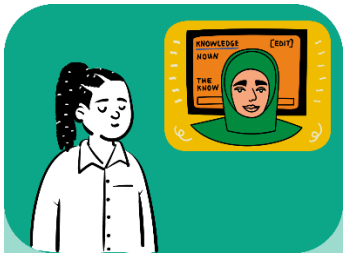
#2 Noor contacted Mobo, whom she met in a session about Wikipedia at a conference in Morocco one year ago. Noor is a 37-year-old Sudanese Archive manager at Cairo University. She wants to get online training to bring her team together to write articles as an activity for a team they formed recently in their free time.



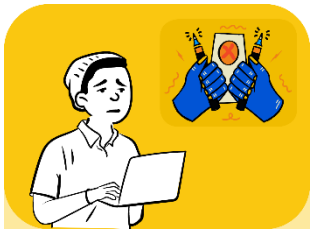
#3 Noor's team has established connections beyond her work. Her team (Tala, Saleh, and Mary) consider themselves a team of knowledge activists. They formed their team after digitising archival data that have historical information about the first 1020 women who were allowed to get university degrees, as men were, at Cairo University, Egypt. Their planned work on open-knowledge articles is voluntary and done during their leisure time.

#4 Noor is personally interested in: (a) developing her team's high-paced collaborative teamwork and digital skills, (b) making a tangible social impact so that she can: 1) Influence large teams of readers through these pioneering Egyptian women, 2) Receive social and professional recognition for raising awareness about the university archive.

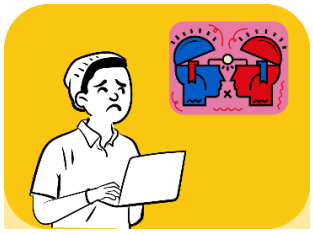
[119] The sketches illustrations from Open Peeps CC0 License. https://www.openpeeps.com/



#5 Even though each person on Noor's team has distinctive personal goals, all members are motivated by the following main goal: after finishing the digitisation process of the archive, they planned to increase awareness about these women and their achievements, so they decided to reach out to Mobo, asking for training on creating new Wikipedia articles. Each member wanted to ensure that those Egyptian women graduates would get the biographical articles they deserved. They also aim to increase women's biographies on Wikipedia.



#6 Mobo has always been frustrated by edit conflicts and the inability to use real-time collaboration in training events, which newcomers are more comfortable with due to the prevalence of real-time coordination and collaboration technology in their day-to-day lives.



#7 Mobo's excitement to experiment with real-time collaboration in his training has led him to attend a session, in 2018, at a Wikimania conference about Real-Time Collaborative Editing. Mobo was thrilled. However, he left the audience with unresolved questions about the social challenges of introducing real-time collaboration to Wikipedia article editing. Consequently, he stopped thinking about it, as he was unaware of ways to introduce an alternative real-time collaboration technology to Wikipedia while at the same time being responsible and respectful of Wikipedia's rich social structure and history.

#8 However, he decided to be the first to try a new research-based tool discussed in a Wikimedia UK blog post to meet Noor's team's needs. The blog post discussed how the Wikipedia community and researchers collaborated on addressing the social and technical challenges of incorporating real-time collaboration in a training session.

#9 Aisha, 33, works at a UK travel agency and has 12 years of experience editing Wikipedia. Aisha is personally interested in: (a) increasing the quality of newcomers' contributions to Wikipedia, (b) ensuring that the Wikipedia community has a sustainable flux of volunteers with a high retention rate through improving the newcomer's experience, (c) meeting others from all around the world and having fun. She read

the same blog article about WikiSync and plans to give it a try by providing support for newcomers and collaborating with them on articles.

#10 The blog post discusses a research-based tool built to provide Wikipedia trainers with all that they need to tailor step-by-step online training for newcomers on Wikipedia's culture, tools, and the new real-time collaborative model for creating Wikipedia articles. It will also provide Mobo's team with an environment with resources to collaborate on learning, researching, and then co-writing Wikipedia articles in real-time.

#11 Once an article is reviewed, it will either be submitted as a draft or to the main Wikipedia with a notice on the article's discussion page asking the Wikipedia community to avoid harsh criticism against the newly trained editors and facilitate a constructive discussion.

The tool notifies the Wikipedian community about the training or article review process, which helps Mobo be transparent about how these articles have been written and allows the community to support Mobo in his training.

The real-time collaboration will be facilitated by an open-source conference tool (Jitsi) and an open-source text editor (Etherpad), which aligns with his values for only using open-source tools.

#12 Mobo noted that setting a target for training would help and encourage everyone in the training to collaborate on achieving it. Luckily, when Mobo designs the training, the tool has a feature that allows Mobo to set the training target and the team to celebrate their achievements.

Mobo goes on to test the tool by initiating the training event process. The flow of how the training will be taking place on WikiSync is illustrated in a later section. However, before this, we need to revisit the WikiSync tool's conceptual model developed through the previous sessions.

The Business Process of the WikiSync System

The business process of the WikiSync System shown in Figure 7.15 was designed to help the participants understand the manner in which the WikiSync interface design was developed.

Following the scenario discussed earlier, Mobo's training has a list of articles to be created. Mobo can publish the event to the WikiSync and Wikipedia communities, which can review the training plan and suggest training material or sources for referencing when writing the articles. Wikipedians can also join the training as participants to support writing the articles or as co-trainers to support the trainer. Another option for the community could be favouriting the training so that they would receive a notification whenever someone requests help to review a section of an article or an article. Only articles that have received a review from the trainer or an experienced editor can be submitted to Wikipedia's main space. Trainees confident with their articles without the review can submit them to the draft space, which could be a replica of MediaWiki on the WikiSync system or on Wikipedia's draft space, where more work, such as Wikifying the text, could take place. The process is an enhanced iteration from the fifth session review activity.

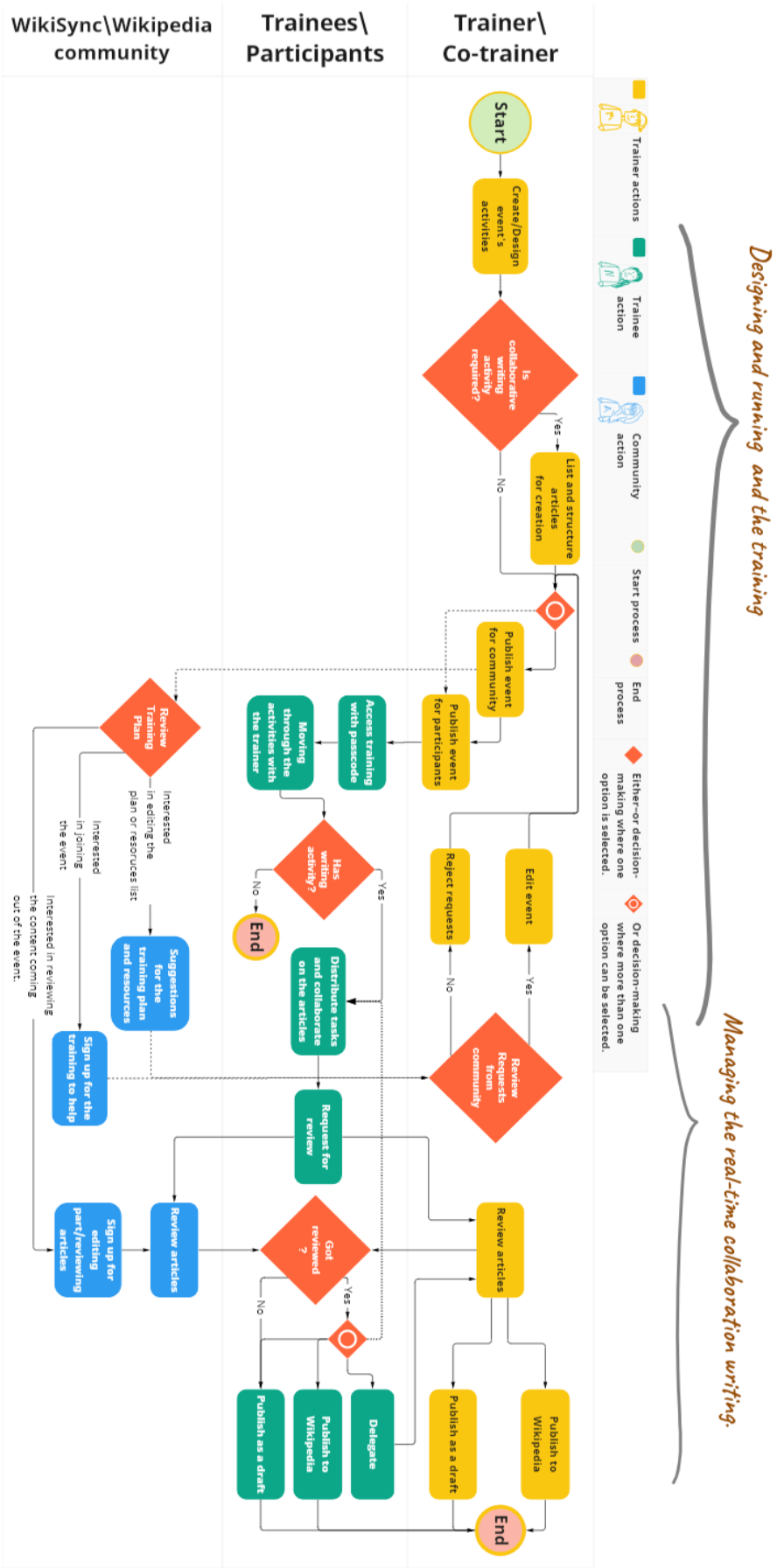


Figure 7.15 WikiSync process.

System Sketches

The following sketches based on the earlier business process were shared with the participants during the fifth session, “Create”. I went through the following fourteen interface sketches to show them what using the system would look like. Some interfaces had multiple variants of the design layout ideas for the fifth session’s participants to decide what was better.

The presented sketches result from multiple iterations and feedback from the research collaborators. As shown in Figure 7.16, I used a Wacom pen computer[120] to sketch the interfaces on the Miro board. This stage required revisiting the latest UX/UI industry practices and Nielsen’s (1994) usability heuristics. The decision for designing low-fidelity was to give the participants the impression that this was not final yet and that any changes would not result in significant effort needing to be invested.

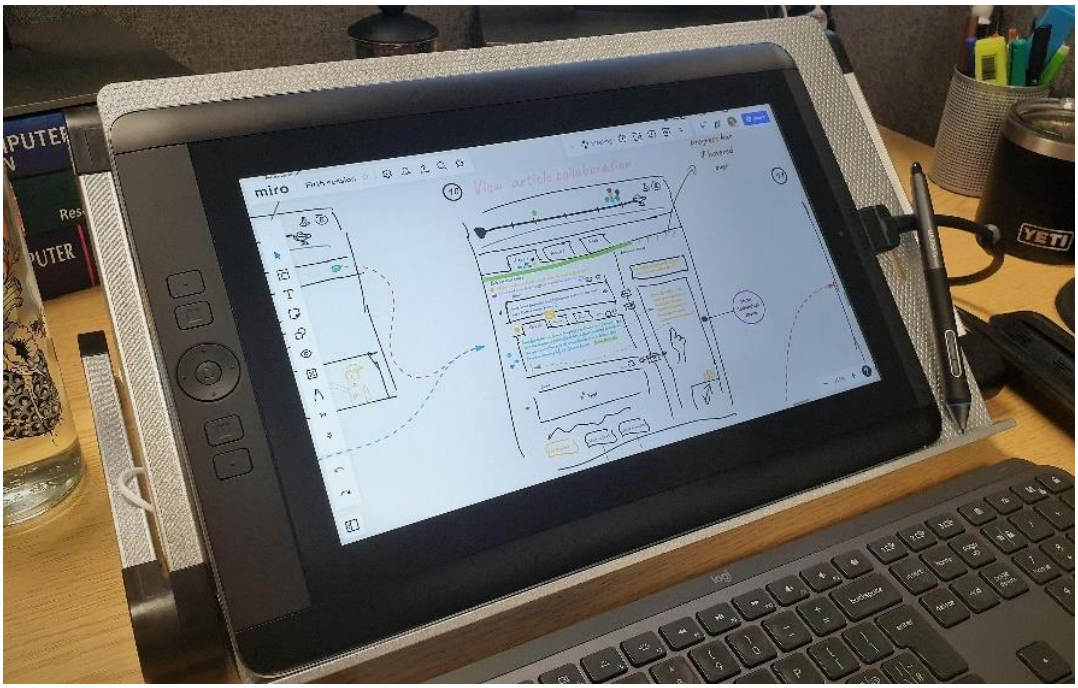


Figure 7.16 Miro and Wacom tablet used for sketching low-fidelity prototype.

Back to Mobo’s training scenario, the next sketches illustrate the process of Mobo, using WikiSync to run a Wikipedia training event with a real-time collaborative writing activity, which will be as follows:

1) We started with his first onboarding to WikiSync. The landing page has the master community interface.

[120] https://www.wacom.com/en-gb/products/pen-computers

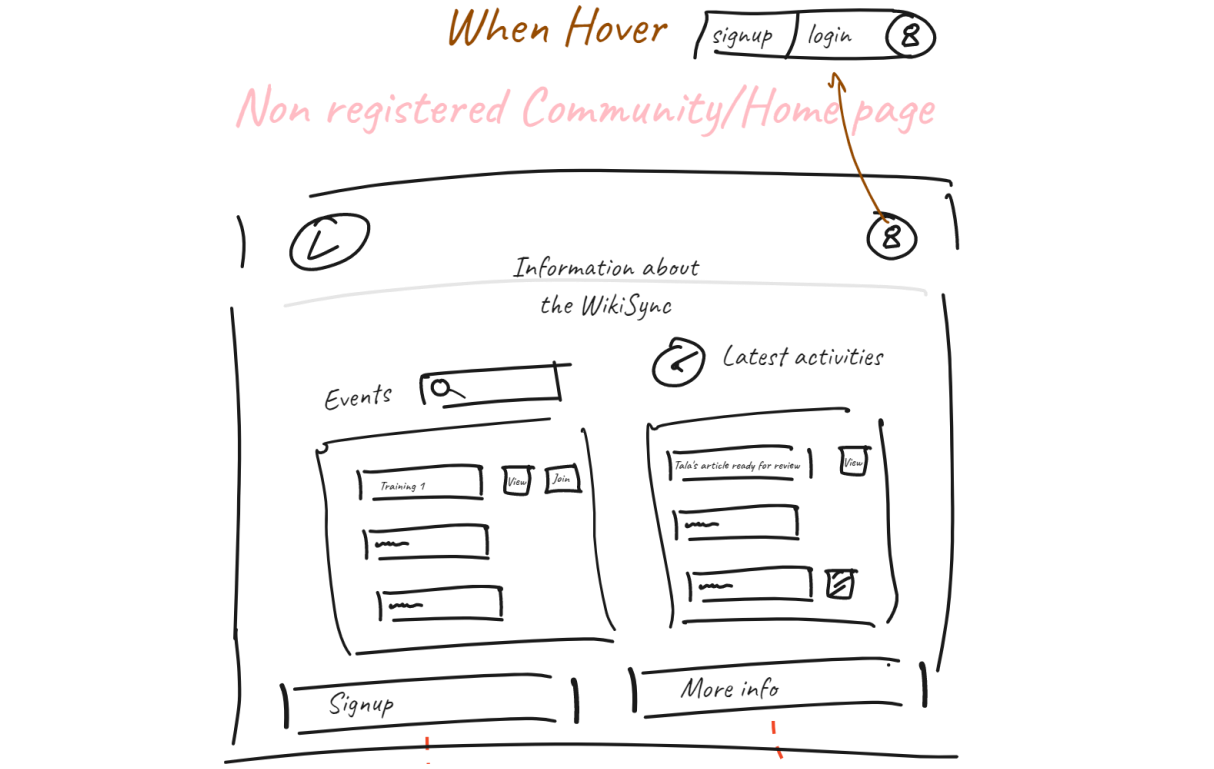


Figure 7.17 Low-fidelity sketch of WikiSync community space interface.

2) The next two interfaces were about learning more about WikiSync and creating an account, as seen in the following figure.

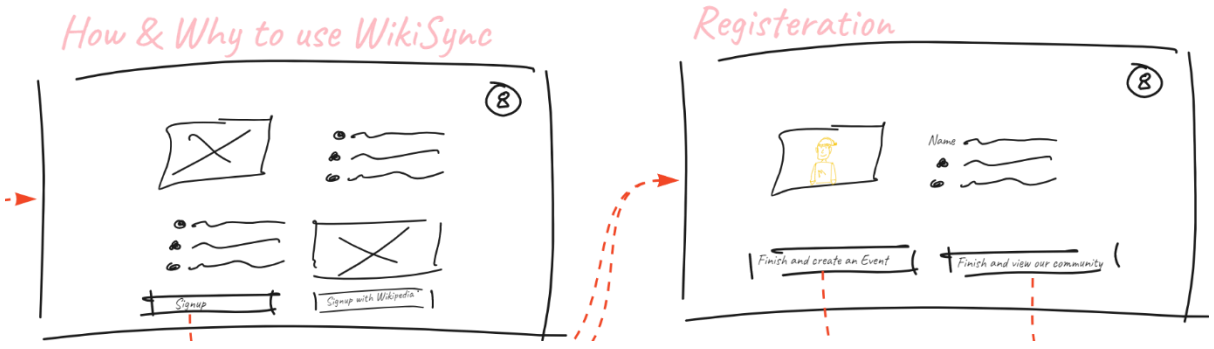


Figure 7.18 WikiSync information and registration interfaces.

3) The following interface is where Mobo designs the training for Noor and her team.

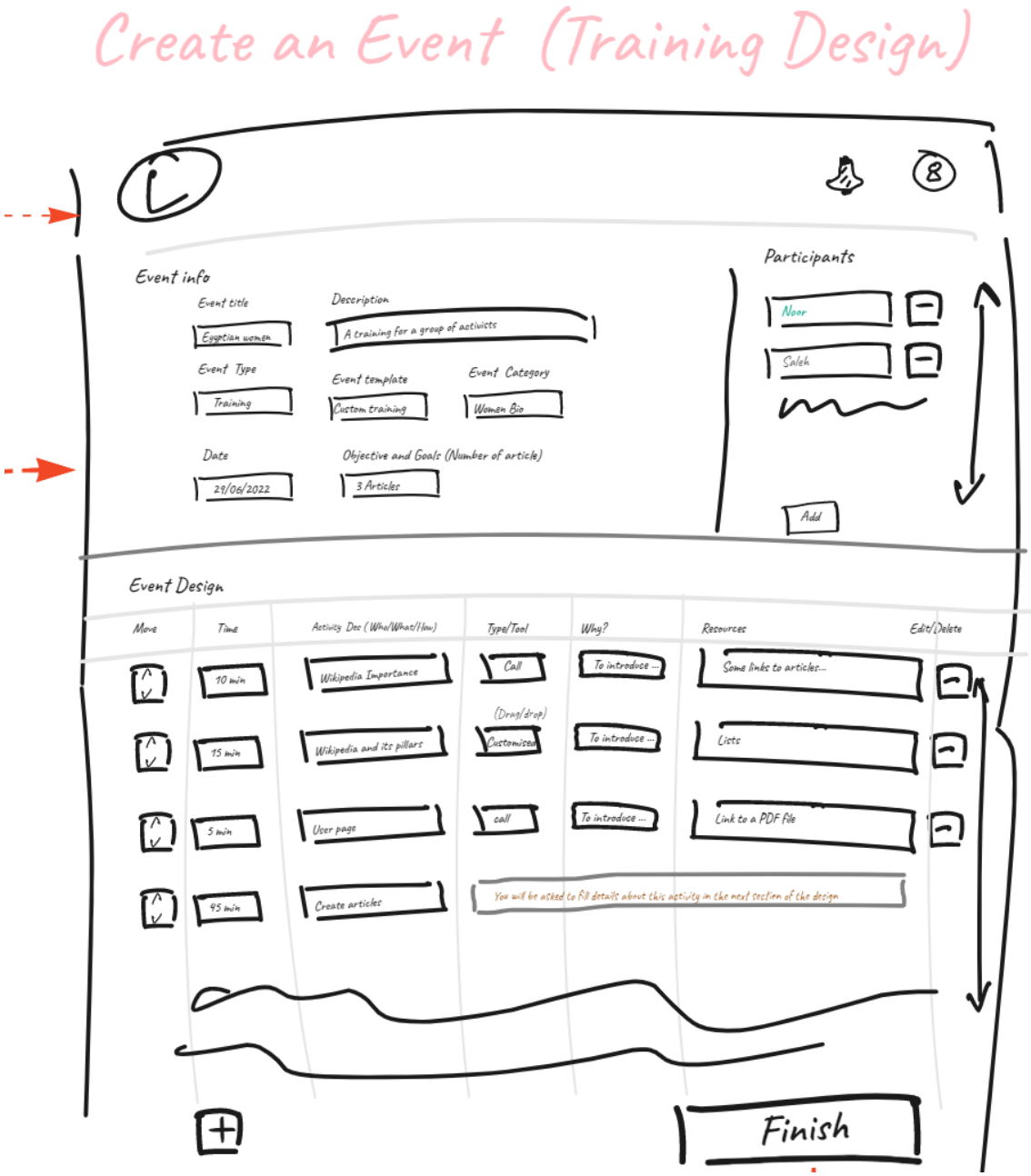


Figure 7.19 WikiSync interface for designing a training session.

4) Since the fourth activity of training includes collaborating on writing articles in real-time, the following interface allows Mobo to add resources for each article, including links to instructions and sources for referencing the articles or event-selecting templates for the article's structure available in the right bar. In addition, the article sections could be set to be written by an individual or open for collaboration.

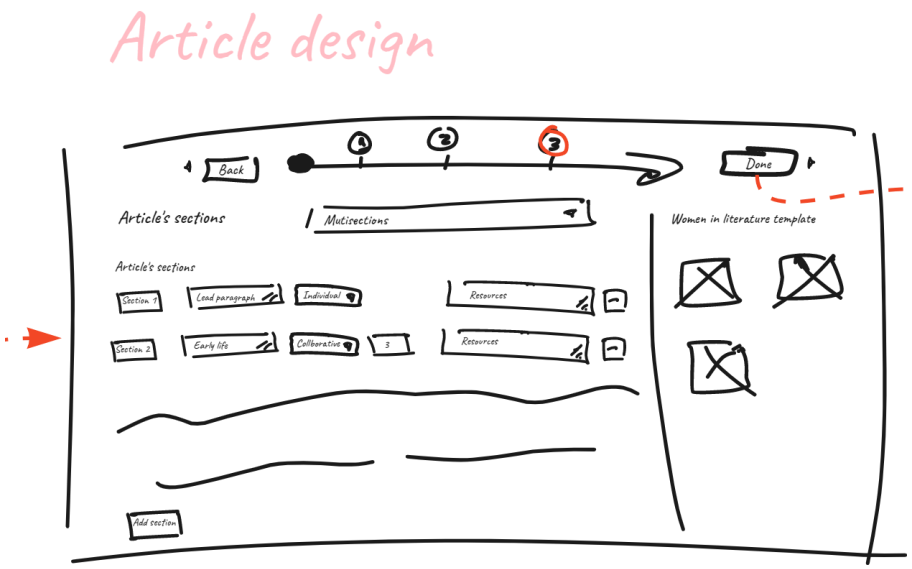


Figure 7.20 WikiSync interface for designing an article plan.

5) Once Mobo is confident with the training design, he can publish the event. If he allows editing, the worklist is activated, and then participants can suggest resources for the articles.

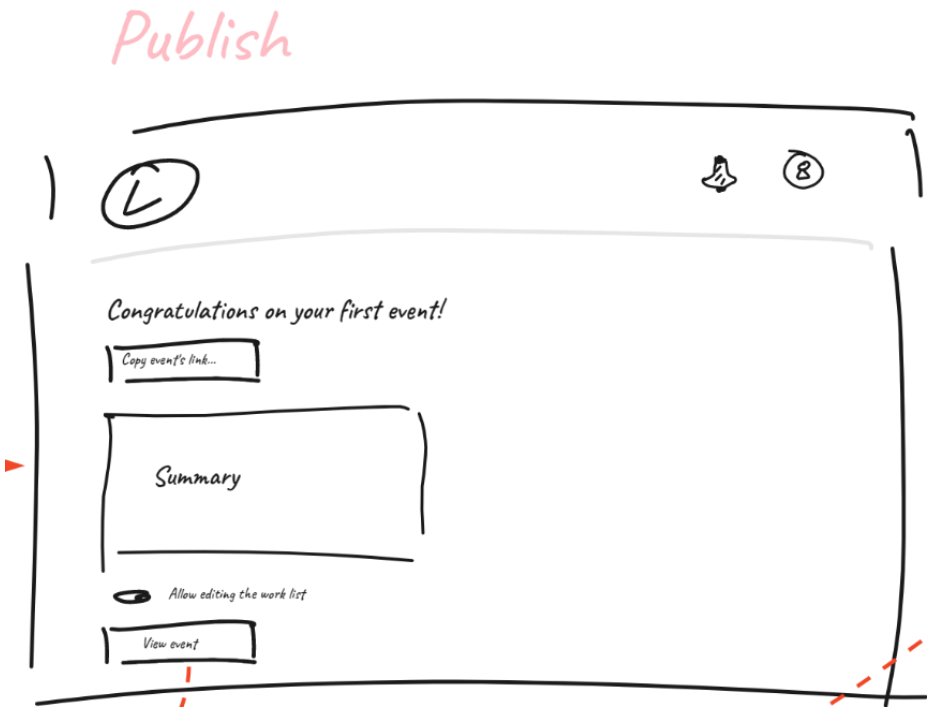


Figure 7.21 WikiSync interface for publishing an event.

6) Clicking on the view event button in the earlier interface will send Mobo to the interface shown below. Aisha from the community can request to join the training event or mark it as a favourite to get notified about articles that need reviewing. All participants can

view the training activity agenda, discuss them, and add resources to the writing activity. All participants have the same view except for the buttons at the end. Participants listed in the event get an access code that allows joining without creating an account by pressing the button in the middle, shown in green. The blue one on the left is for those who signed up as community members offering help and must have a Wikipedia account registered in WikiSync. This separation helps Mobo prevent vandalism or abuse and control the community's access to activities (viewing, reviewing) or communication channels. The button in orange on the right is only visible to Mobo to start and join the event, which sends him to the following interface in Figure 7.23.

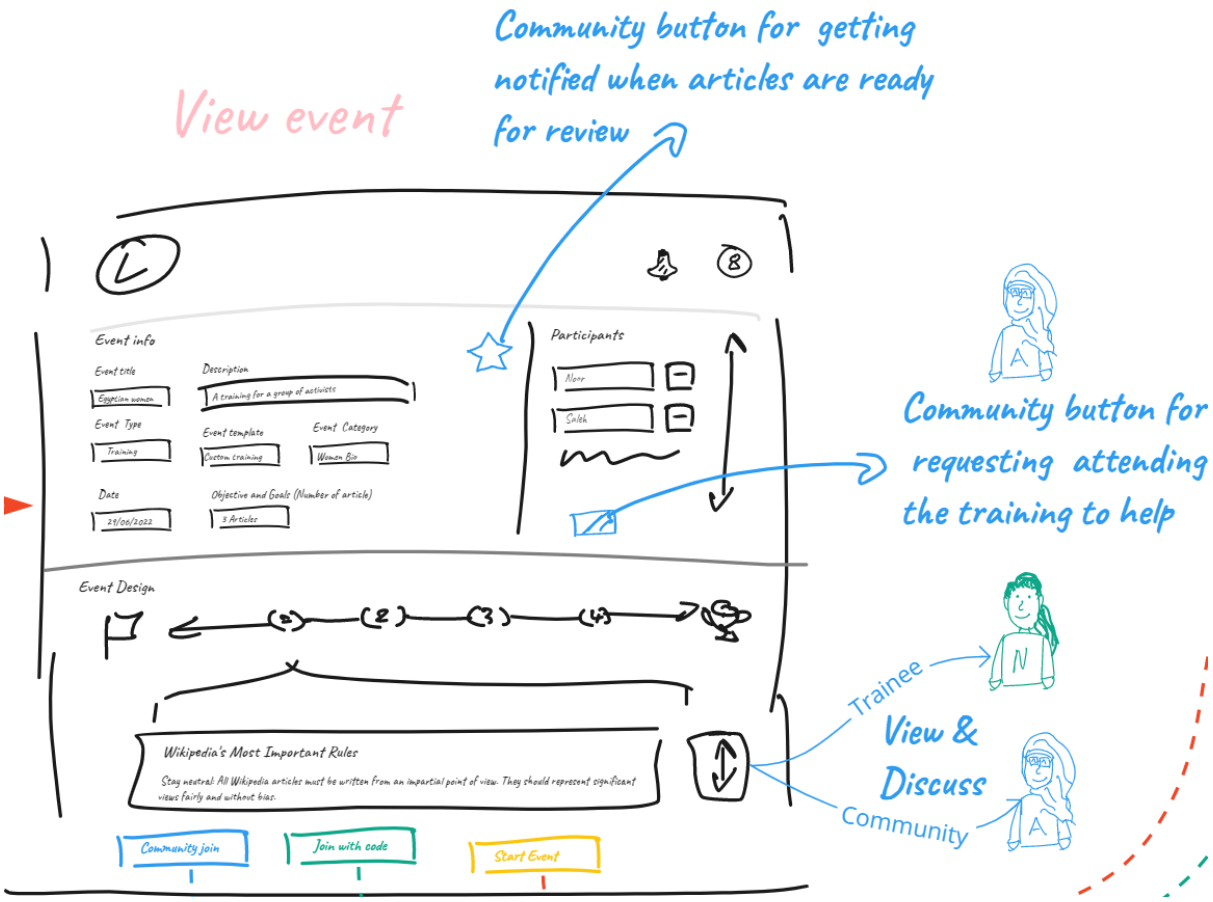


Figure 7.22 WikiSync interface for viewing an event.

7) In the following interface sketch, Mobo is shown going through the training activities one by one. Everyone in training can see the progress by hovering on the progress bar. The progress in Figure 7.21's interface shows participants in an activity where Mobo demonstrates an article's building blocks to his trainees. To put things into context, I included Figure 7.24, which uses an actual conference call interface, showing how the training call would be embedded inside WikiSync. The following section covers the

interface used for writing the articles list prepared by Mobo, which was reviewed in detail by the “Create” session participants.

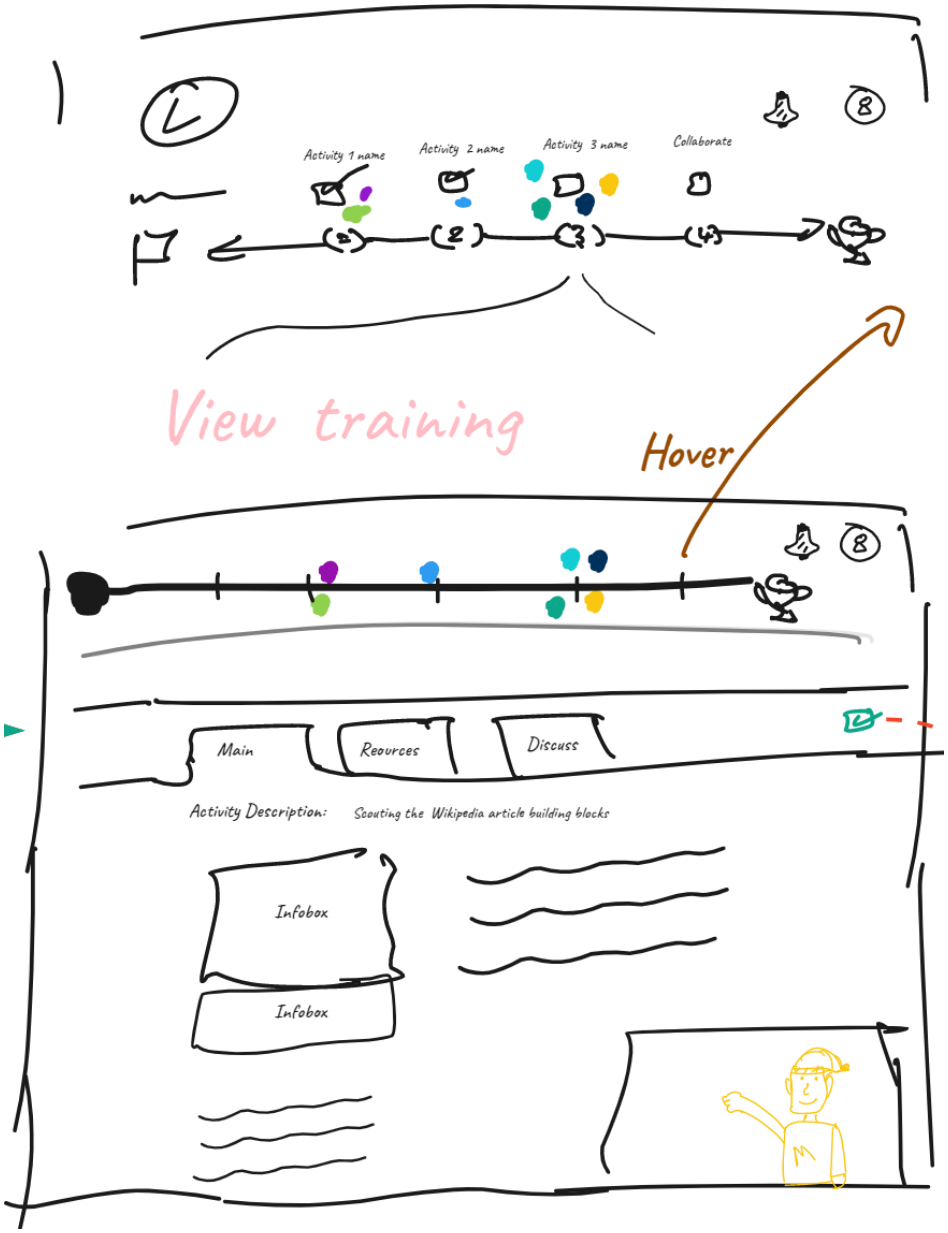


Figure 7.23 WikiSync interface of an ongoing training.

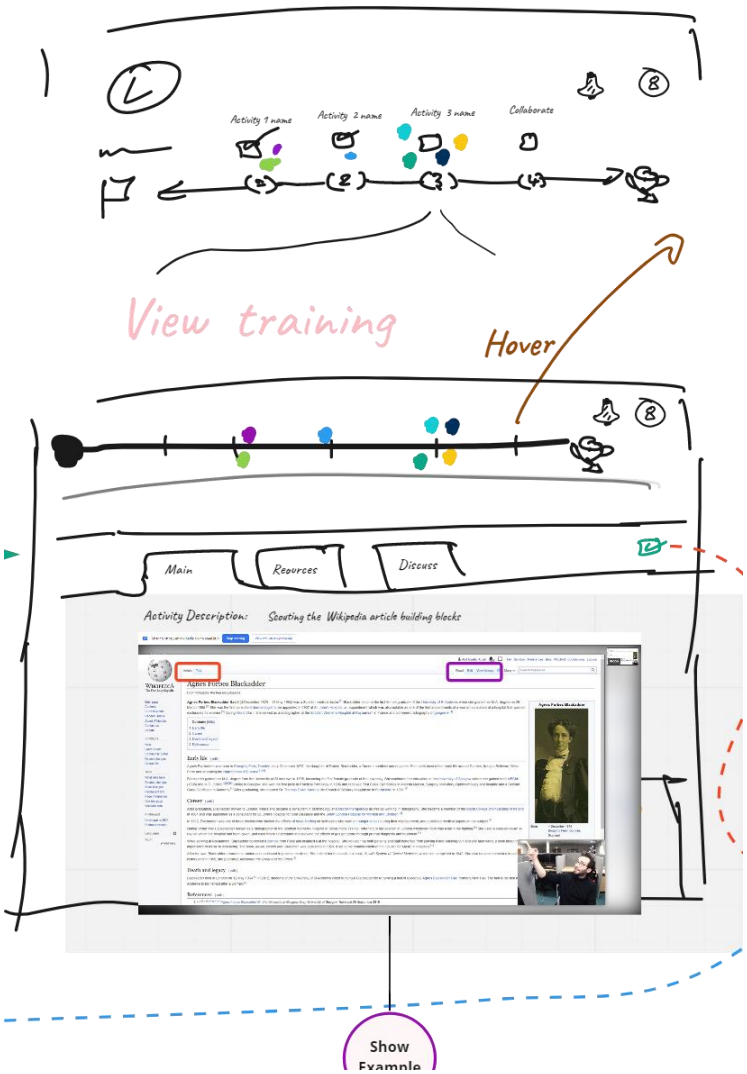


Figure 7.24 WikiSync interface of ongoing training that uses Jitsi conference call.

8) The interface in Figure 7.25 is shown to participants ready for the collaborative editing activity, who would see the articles to be created in several tabs, with a progress bar for each article below. Mobo planned some sections for collaboration and some for individual work. Once a trainee or trainer clicks on a section, the text and tabs for resources and discussion appear. The **'Discuss'** tab can be used as a space to comment or a conference call for this section. The community can see and edit the text if participants check the review checkbox. Any participant can work in a private section after they indicate the title of what would be worked on. Others can see from a pace under the progress bar the title of what is being developed in participants' private spaces.

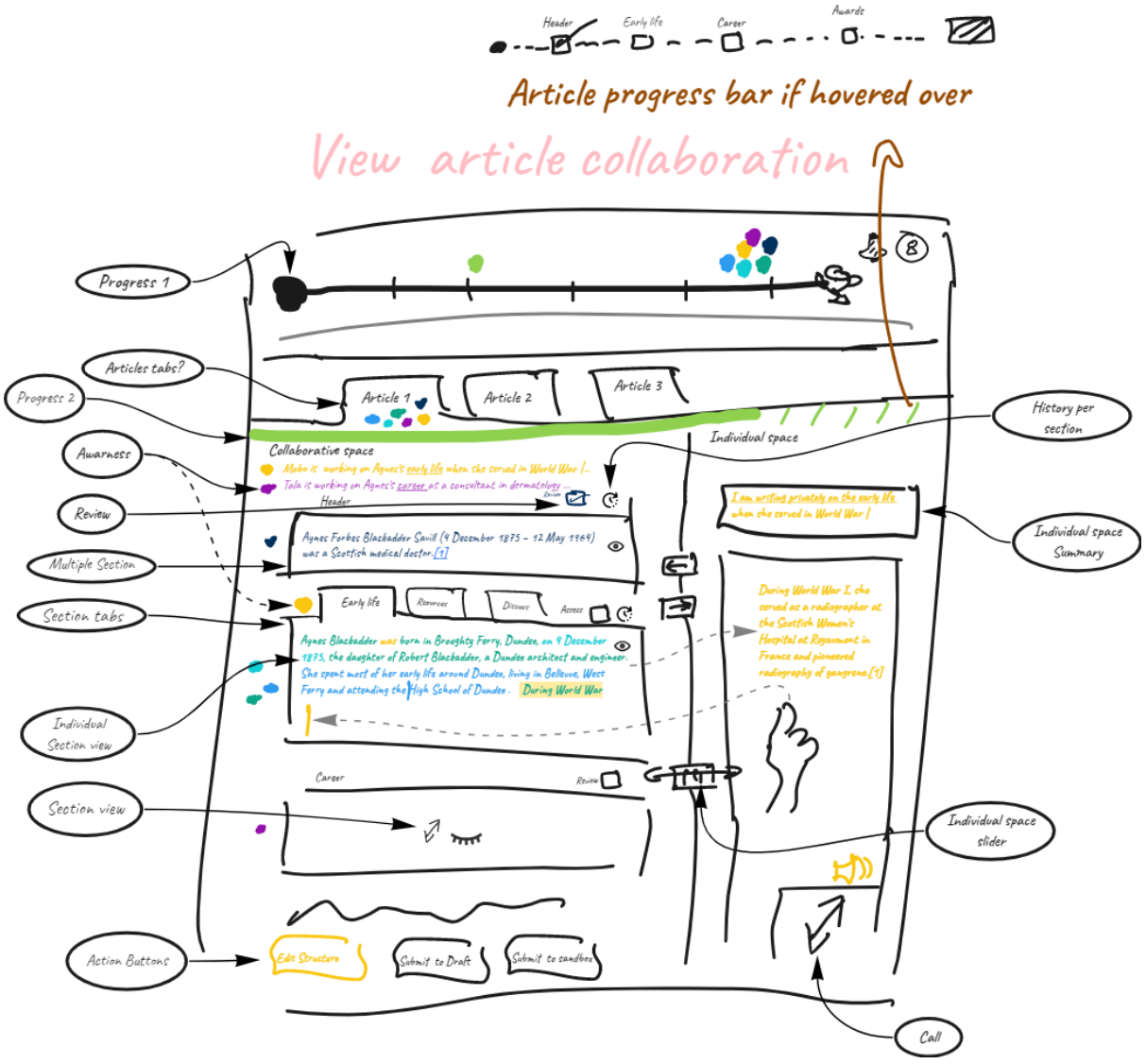


Figure 7.25 The first proposed interface for WikiSync real-time collaborative writing space.

9) Once the article is ready, the trainer and trainees can move it to the sandbox or a draft. Wikifying could happen in the writing process, the sandbox, or the draft space.

7.3 WikiSync Design Cycles

As a recap, this research framework is divided into several phases, where the Discovery phase has the "Identify" and "Define" sessions that delve into the community challenges. The Conceptualisation phase defines the user groups, the usage scenario, and guidelines for a new system. Finally, the Design and Social Voting phases iterate on and refine the system concept leading to the co-design of a prototype for the new system. This phase consisted of a "**Create**" session, and two iteration sessions that mostly focused on the collaborative space interface. Seven participants took part in these sessions. Almost the same participants took part in these sessions. The session activities design was mainly covered in 7.2, where participants walked

through the sketches that resulted from the AT. Participants' activities involved sketching and iterating on the design concepts and ideas.

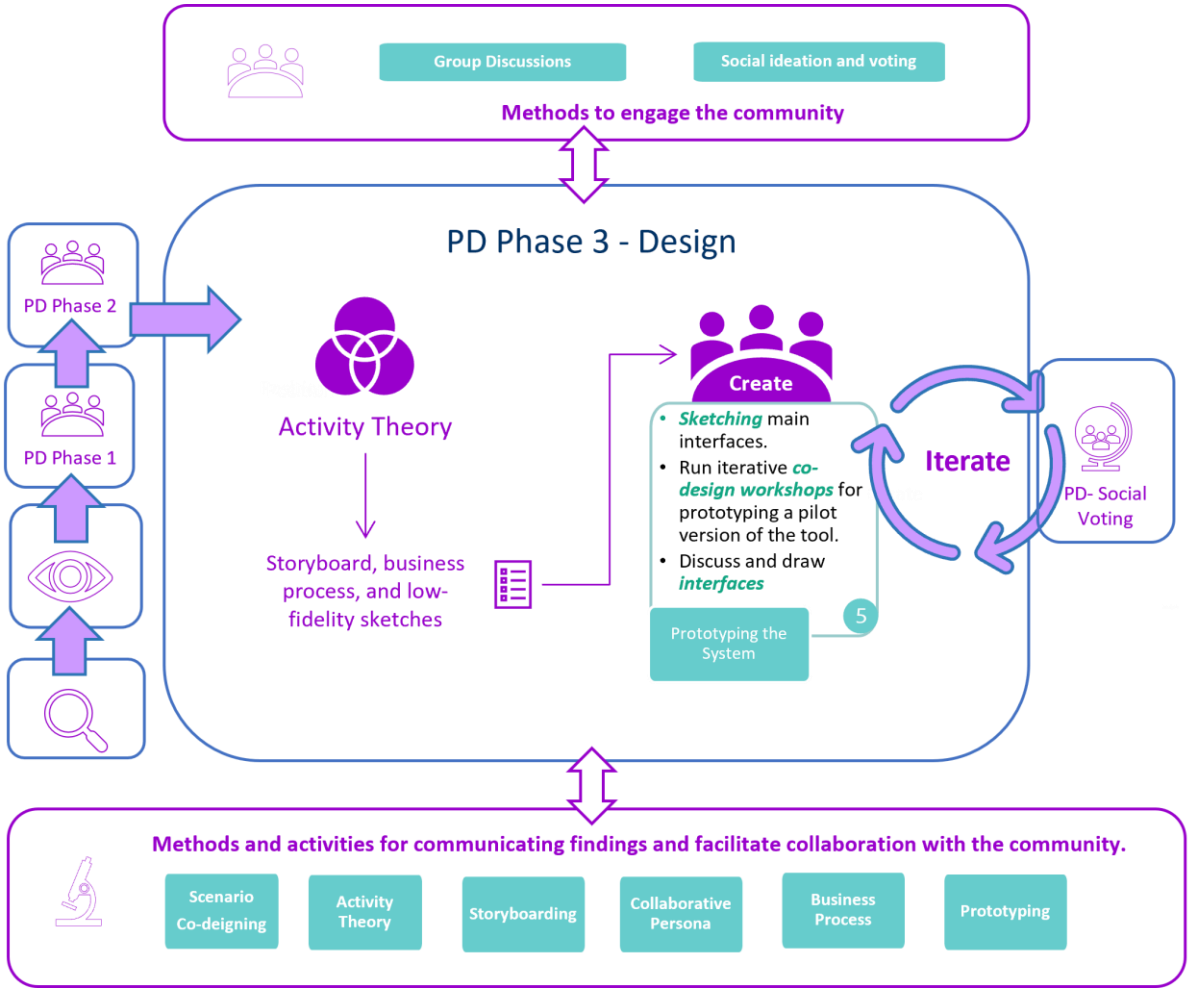


Figure 7.26 The Design phase and its methods.

The "Create" session and its iteration used the Miro board shown in Figure 7.27 with some alterations to fit the thesis page size. The original board put the following series of board frames in a horizontal sequence. Frames 1,2,3, and 5, shown in Figure 7.27 below, have already been discussed in the previous sections.

1) Previous sessions' frames covered the WikiSync user groups, collaborative persona, and individual characteristics of the scenario's trainees, such as their background and previous Wikipedia experience.
2) WikiSync scenario.
3) WikiSync business process.
4) "Our task for the day" (discussed below).

5) The low-fidelity sketches with spaces for feedback on the design from three perspectives (trainees, trainer, and community members)

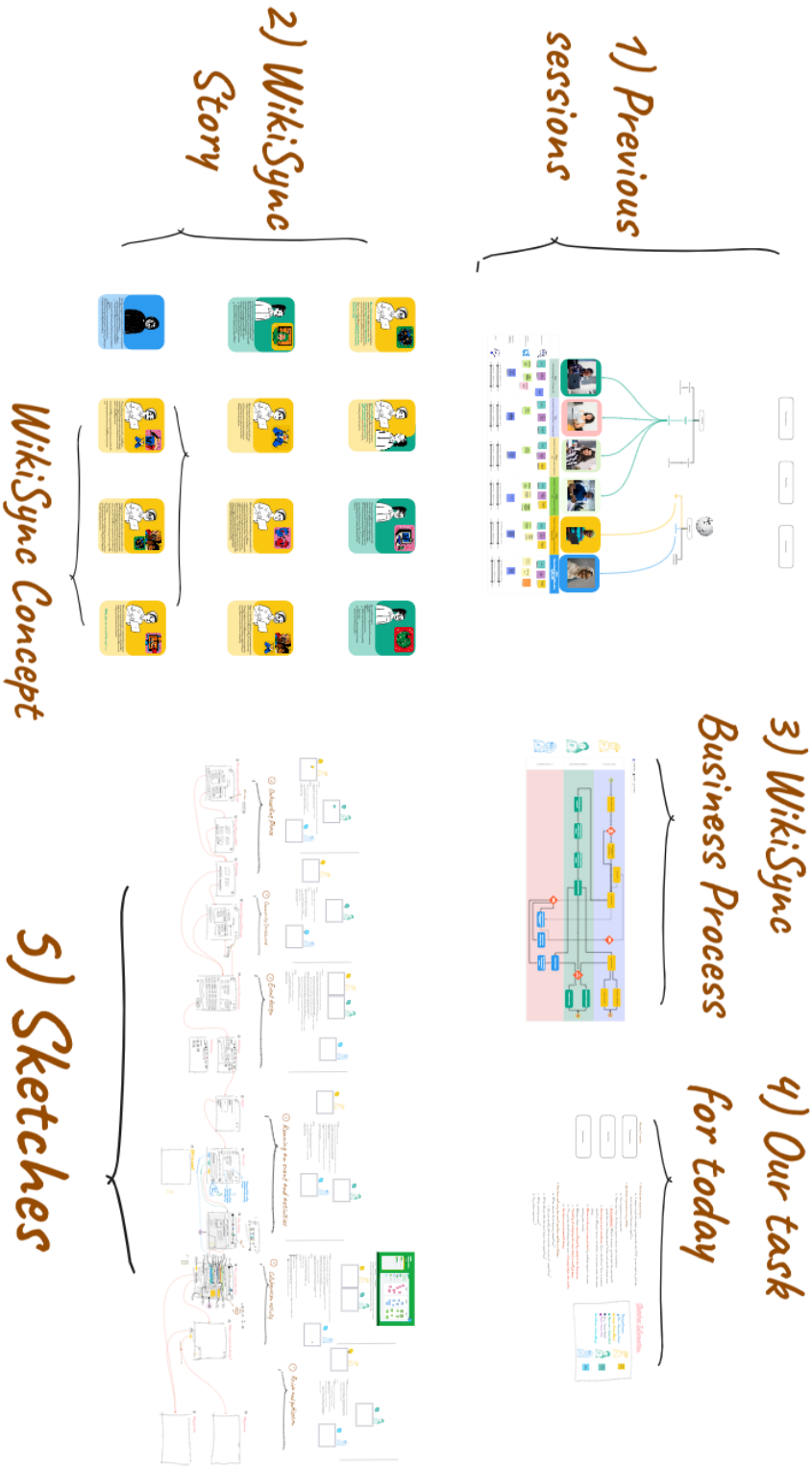


Figure 7.27 Fifth session Miro board updated to fit the thesis page size.

There is a growing body of literature on participants' capabilities to produce and innovate (Hyysalo et al., 2016). However, the time limitation of the sessions and the diverse technical backgrounds of the participants have led me to clearly define and set the expectation for their tasks in the session by introducing Frame 4, shown in the figure. The frame listed some themes from the thematic analysis discussed in Chapter 6, such as transparency and flexibility and then the following three points:

1) Setting expectations. This item repeats a point raised at the "Identify" session on being realistic about what could be achieved during this research. I have explained that, even though the whole system is sketched to get a feel for the flow of activities in the system, one must focus on the most important features to benefit from the opportunity of introducing real-time collaboration in training. This is mostly the case in the "Create" iteration session that discusses the real-time collaboration interface (see Figure 7.25).
2) "No design experience; no problem." This ensures the participants that, even though they are not UX/UI experts, their experience in Wikipedia is essential. It also covers some tips, adapted from Nielsen's (1994) usability heuristics, on what to look for when reviewing the sketches.
3) They were told that the focus could vary when specific interfaces are reviewed and that they generally should be prepared to be asked the following questions:
    a) What can we do to make improvements?
    b) Can you quickly sketch a better version of something?
    c) What did you like the most about it?
    d) What could go wrong?
    e) Any other comments?

Then, they were introduced to the diagrams' legend, and the sketches were reviewed.

This stage involved sketching alternative interfaces presented by the participants using pen and paper or the session's Miro board (see Figure 7.28 and Figure 7.29).

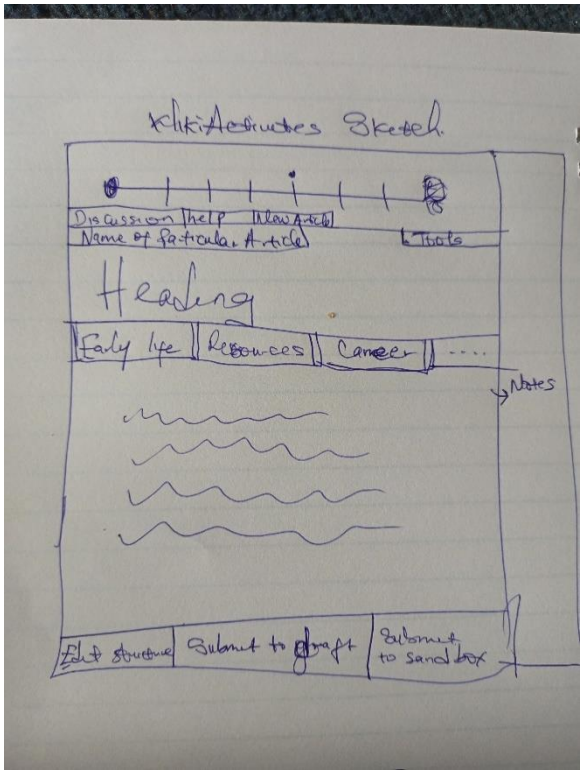


Figure 7.28 A participant pen and paper sketch to improve the real-time collaboration interface.

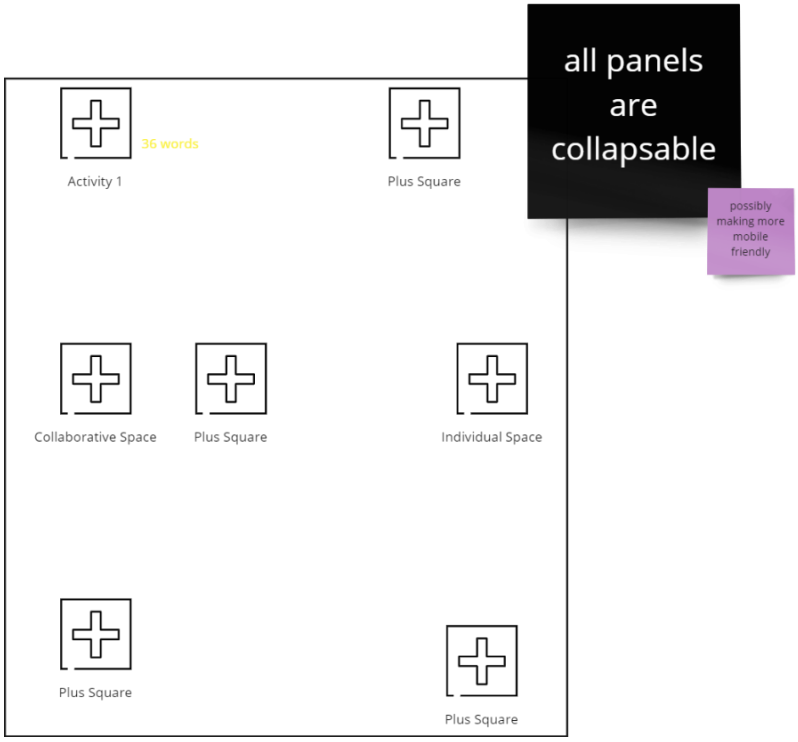


Figure 7.29 A participant sketch using the session's Miro board to improve the real-time collaboration interface.

Section 7.3.1 covers the highlights from the fifth PD "Create" session and its iteration.7.3.2 covers the Social Voting *phase,* where the broader community is consulted on the ideas that have emerged from the study. Lastly, Section 7.3.3 discusses how this iterative process has

helped inspire the high-fidelity prototype of WikiSync, which was demonstrated in the last session for discussing the future of WikiSync.

7.3.1 "Create" Session and Its Iterations

After demonstrating the sketches, the participants praised many features that matched what they expected (such as the training tailoring and private writing space). However, this thesis focuses on the challenges of such a scenario and how the iterative process and other factors, such as mutual learning, have helped advance the design and bring impetus to the innovation process. Therefore, the discussions cover such questions as what could go wrong? What can be done to improve something? What did you like the most about something? And opening the discussion for any other comments or suggestions. Thus, the main concerns that emerged from the discussion points are summarised as follows:

The level of transparency. There was a concern that if the tool produced articles that went directly to Wikipedia's main space rather than that draft space, the Wikipedia community could be critical of the transparency level at which the article is produced. For example, suppose a group of editors wrote an article with a promotional tone in WikiSync and prepared plans to challenge the deletion of a promotional article. The Wikipedia community could be overwhelmed by the systematic debate against its rules. This problem is not unique to WikiSync and could happen within Wikipedia or any other external tool, and it was raised to protect WikiSync against potential stumbling blocks.

The question is, will the tool allow Wikipedia community members to access the full training or only the editing part? What are the risks or benefits of allowing such access? How to control it?

This discussion defined that a limited, conditional access community to the editing part of training could help increase transparency and support the trainees.

Authorship. A concern was raised about who gets the attribution on Wikipedia for the articles produced in WikiSync.

It was agreed that whoever comes from this real-time collaboration in WikiSync to Wikipedia should be given attribution in the edit summary. Therefore, either one participant creates an article in Wikipedia's draft space, and each trainee submits a section she/he has contributed to the most, or one trainee or trainer submits the whole article and gives credit to others in the edit summary.

Also, the feedback from sessions stresses the importance that, when working on the articles, WikiSync participants should realise that their contribution is under a compatible license to Wikipedia, such as CC by SA.

Finding a balance among desirability, feasibility, and viability of potential solutions. My research aims at producing a design that programmers could implement with a reasonable budget. Therefore, the first version of WikiSync design must be desirable to the community, realistic from a technical perspective, and sustainable, serving Wikipedia's vision.

In the fifth session, we had to be careful with our requested features. We started with the why of any proposed feature and declined (or postponed) complex features with no prospect of implementation in a reasonable timeframe. Also, we rejected oversimplifying the technical implementation by stripping essential elements that could compromise the new collaborative process activity, leading to a breakdown. Below are two examples of such debates.

A. **Wikitext.** Adding a full Wiki editor and syntax in real-time could increase the tool's technical complexity and overwhelm trainees. Thus, we agreed on making the WikiSync interface as simple as possible, focusing on writing the article at this stage. Wikifying the text can take place at a later stage:
    a. On Wikipedia's draft space, or
    b. On a space inside the tool with an identical version of Wikipedia's MediaWiki software hosted by WikiSync. Having it inside WikiSync could provide the trainees who would take on the wikifying task a safe space to experiment and develop the articles created after the real-time collaboration and grow confidence in shaping articles ready to be shipped to Wikipedia's draft or main space.

The hope is that future versions of WikiSync would develop to produce a real-time collaborative article with Wiki syntax ready to be published directly on Wikipedia. Still, the technical feasibility is a critical challenge here and one of the reasons that has led to stalling in Wikimedia's introduction of this feature.

However, one of the participants highlighted the importance of keeping track of citations. At least, the source of information needs to be shown, as they could be lost in the process if two separate people were to work on different stages of the article.

This has led to adding a basic text styling bar at the last design prototype.

B. **Mobile compatibility** is another example where a balance had to be struck. Having a design friendly for mobile devices is very important and was stressed by many participants. Still, we had to remind ourselves that this was another complex issue Wikipedia faces and

goes beyond this project's focus. Since it was not directly connected to the research focus of introducing real-time collaboration, we had to move on and defer the discussion on the implementation of a mobile-friendly interface.

Nevertheless, this push from participants on mobile-friendly design has led to prioritising design decisions that are more menu-driven user interfaces. This could be the basis for future work on a mobile app for WikiSync.

Iteration of the "Create" session.

Since this session is an iteration of the fifth one, discussion points are similar, focusing on the real-time collaboration interface in Figure 7.25. Again, the earlier discussion points are picked up on, with some participants doing new sketches of the interface, and new concerns and solutions have emerged.

In addition to the positive feedback, the main concerns and discussion points are the following:

User account. As shown in Figure 7.22, the interface design provides access with a passcode, and there is no need to log in to minimise the friction and tasks newcomers should do, i.e., creating separate accounts for WikiSync and Wikipedia. This decision is based on earlier feedback from the research collaborators at Wikimedia UK. However, one participant suggested having a single-user login where participants can log in using their Wikipedia accounts, which could simplify the attribution process. As with Wikipedia, those who prefer not to log in with their accounts to WikiSync, their attributions should be connected to their IP address and asked to license their contributions under CC by SA. Such adoption could help avoid getting "bad faith actors [...] who the foundation and the community spend a lot of time [filtering out]"—Participant A. This addresses the concern of using WikiSync as a security bypass for that small number of bad actors the community is trying to keep away from the site.

IP Blocking. One of the participants worried that Wikipedia's user filter systems and IP blocking policy would create problems for WikiSync users setting up from the same or similar IP address. To prevent this issue, one participant suggested that trainees create accounts on multiple Wikimedia projects, such as WikiCommons, and Wikiversity, for users to edit Wikipedia. Also, having a Wikipedian belonging to a user group, such as an Event Coordinator[121], capable of confirming the accounts[122] would be another way to avoid this type of blocking.

[121] https://en.wikipedia.org/wiki/Wikipedia:Event_coordinator
[122] https://en.wikipedia.org/wiki/Wikipedia:User_access_levels#Confirmed_users

Another case is many trainees publishing from the same or similar IP address. A participant indicated that WikiSync real-time collaboration has the advantage since one person is publishing the group's work and that there would be no instance where more than one trainee would hit the publish button simultaneously (something that Wikipedia could interpret as vandalism).

Complexity. The interface in Figure 7.25 illustrates becoming dense with information. It is described as "a comprehensive interface" that saves accessing many systems to conduct training but "possibly trying to do too much." This discussion is later connected to mobile compatibility when a participant designs an alternative interface with an "all panels are collapsible" version of the interface to reduce complexity and make it mobile-friendly.

This, combined with attempts by other participants to sketch something less complex, has shaped the final prototype (see Figure 7.28 and Figure 7.29) with reduced complexity compared to Figure 7.25's interface.

Fallback channels. Following on from the complexity concern, one participant noted a "great potential for the application to break down." This connects to earlier discussions about being careful with how "bandwidth hungry" the tool is and the need to accommodate those with Internet access that may not support the conference call.

This stresses the importance of providing the trainer with optional settings that include an asynchronous communication channel for collaborators who are incapable of joining the real-time collaboration method and its call.

Duplication of effort. One participant raised concern about the design flaw in Figure 7.25: Suppose someone is working on one section without viewing how other sections are developing. In that case, this might lead to the duplication of efforts, where participants cover the same ideas in different sections. This point has led other participants to note that this also may lead people to link keywords to other Wikipedia pages multiple times in different article sections rather than just on the first occasion.

This has brought on the introduction of the **view full article** feature in the last prototype, which allows a read-only view, allowing reading other sections while editing one of them. Another participant suggested that avoiding multiple links to another Wikipedia page could be solved by another feature notifying the writer that it is already linked or viewing it in the full article view.

Another participant recommended a discussion space for coordinating work on an article level. This was incorporated into the final design, similar to Wikipedia article pages. This could give a newcomer an experience in WikiSync akin to that of a Wikipedia encounter.

Draft space duplication concern. One participant mentioned that the trainer or the system should check that the article about to be created does not already exist in Wikipedia's draft space.

The importance of maintaining a high level of security to prevent real-time vandalism is mentioned again in the iteration session but not discussed in detail due to time limitations.

Other ideas. One participant recognised a great opportunity in the feature discussed in Figure 7.20 interface, which shows "different customised article templates to enable editors to create different articles using the Wikipedia accepted article template". Therefore, how such templates could help with other social challenges and make the training design easier should be further investigated. For example, this may lead to design templates being crowdsourced and verified by the Wikipedia community to be used in WikiSync article writing.

7.3.2 Social Voting Phase

On the 7th of November, a questionnaire was advertised to many Wikipedia communities (see Figure 7.30).

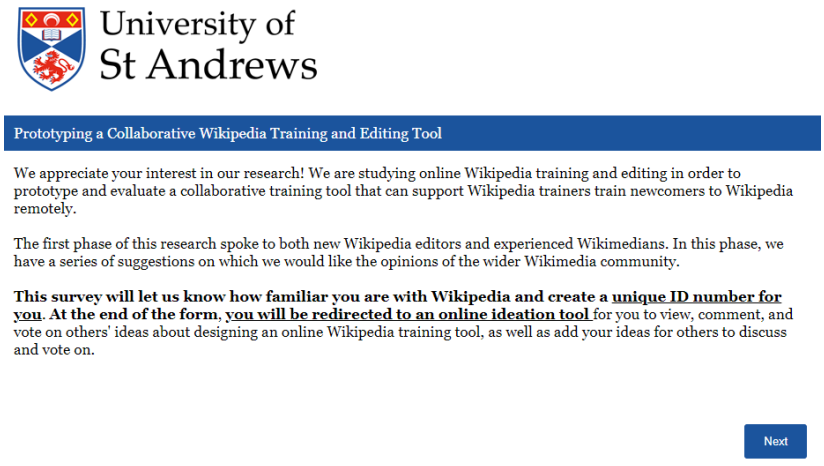


Figure 7.30 A questionnaire advertised for the Wikipedia community to get their feedback on the research outcome.

It enquired about a respondent's familiarity with Wikipedia, giving her/him an anonymous, random ID number, which was used to participate in a social ideation and voting tool called Tricider. 19 ideas were listed on Tricider. Figure 7.31 shows part of the tool displaying two ideas, their arguments, and votes that were received by the random IDs.

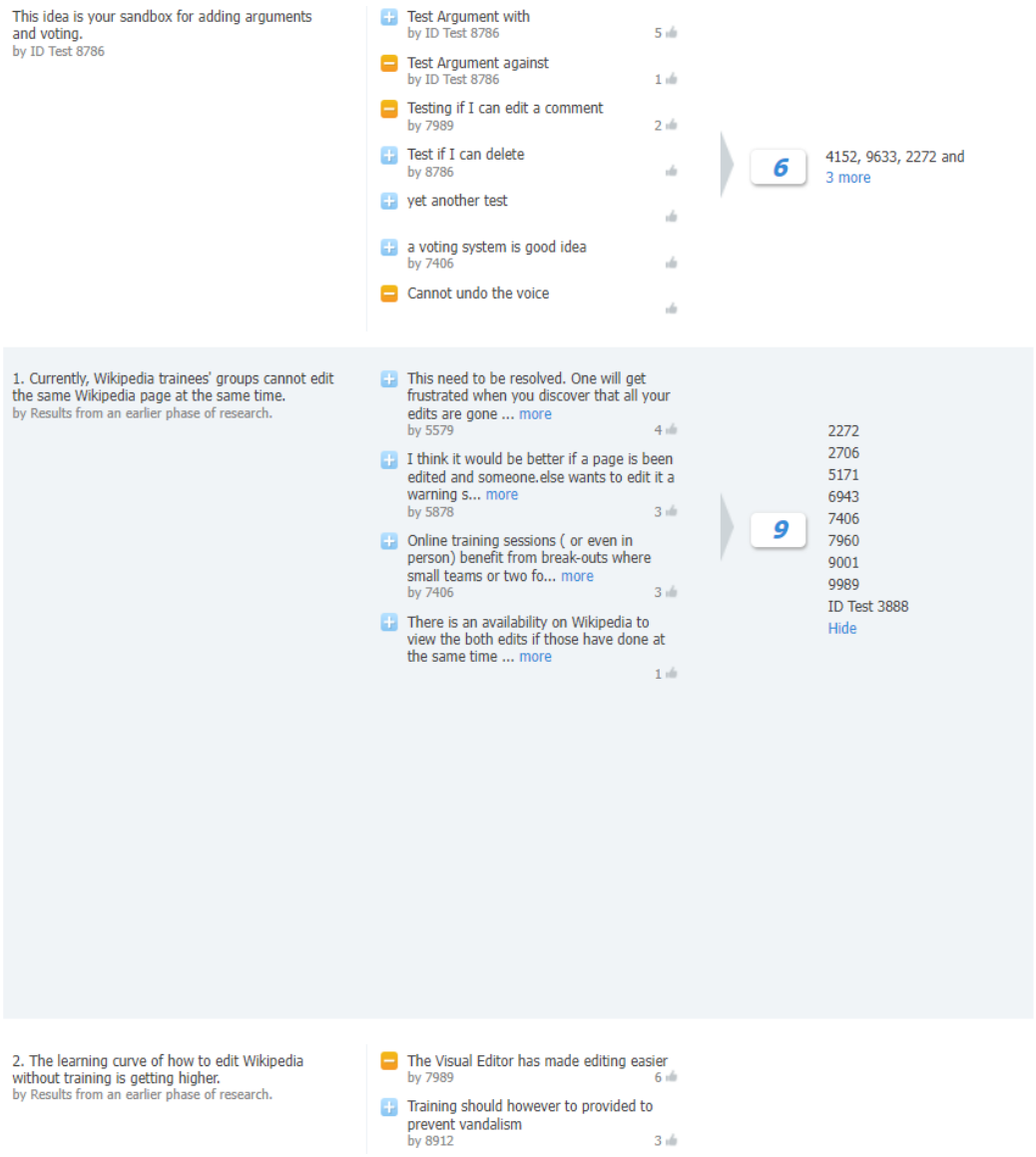


Figure 7.31 A Screenshot of the Tricider tool used in this research.

The full list below summarises all of the sessions' results for the community, including those of the review and iteration. In addition, ideas 6, 10, and 11 had links to interface sketches in Figures 5, 8, and 11 and a pen and paper sketch by a participant in the fifth iteration session.

1) Currently, Wikipedia trainee groups cannot edit the same Wikipedia page at the same time.
2) The learning curve of how to edit Wikipedia without training is getting higher.
3) Recent research shows that running Edit-a-Thons contributes to easing the journey of newcomers, as well as fighting the systemic bias on Wikipedia.
4) We need to invest in improving the trainees' experience in Edit-a-Thons, and Wikipedia training.

5) It is viable to add a real-time collaboration model for creating new articles during Edit-a-Thons or Wikipedia training sessions.
6) We need a new system that allows trainers to design flexible and transparent wiki workshops to train small groups on real-time collaborative Wikipedia writing, starting from the training process and ending with sustainable volunteer engagement. For more info, please view one of the interface sketches for creating event. https://standrews.eu.qualtrics.com/ControlPanel/Graphic.php?IM=IM_6te6VrcnRjX2bxc
7) The system should focus at the start on groups with similar interests and skill levels, such as groups of work colleagues who work at a library archive and care about raising public awareness about the topics they care about.
8) People are getting more familiar with real-time collaboration when working in teams, so it would be helpful to offer such a collaboration option when creating new articles.
9) It is important to have a new system for Wikipedia training built to provide Wikipedia trainers with all that they need to tailor and deliver online training. It should offer an environment for groups of trainees to collaborate on learning, researching, and co-writing Wikipedia articles in real-time.
10) A Wikipedia community member can access real-time collaboration as a viewer or commentator by default and as an editor, should they have permission. For more info, please view one of the interface sketches for viewing event. https://standrews.eu.qualtrics.com/ControlPanel/Graphic.php?IM=IM_bJHUauYyQw16blk
11) A trainer can give a Wikipedia community member access to the real-time collaboration process as a contributor. For more info, please view one of the interface sketches for the collaborative editing activity at the event: https://standrews.eu.qualtrics.com/ControlPanel/Graphic.php?IM=IM_5goiV366sN6FDHU Also, please view, one of the first phase research participants' rough, simple sketches of an alternative collaborative writing interface for the tool: https://standrews.eu.qualtrics.com/ControlPanel/Graphic.php?IM=IM_9tXgnJoV5ivKQ1E
12) Suppose a verified community member or the trainer reviewed the synchronously created article draft. This qualifies the article to be moved to the main space with a summary note on its talk page about the training and the list of people who have contributed to it.

13) In the collaborative writing space, the trainee should have a personal editing space, not shared with anyone. Others should know what and where the trainee is working to avoid duplication. Once confident with their edits, they can drag the text and drop it in its dedicated place.
14) In the collaborative writing space, all panels are collapsible, possibly making it more mobile-friendly.
15) The system should be open-source, facilitated by open-source tools, such as conferencing (Jitsi) and an open-source text editor (Etherpad), to keep the possibility of future incorporation into Wikipedia.
16) When collaborating in real-time on an article, having a conference call accessible by the Wikipedia community as viewers could raise transparency.
17) The tool should make it easy for trainees to get to community spaces on Wikipedia, such as the Wikipedia: Village Pump or the equivalent community-managed page in another language.
18) It is important to involve the wider Wikipedia community through the Tricider tool to collect ideas, discuss, and vote on a new system.
19) Shall we call such a system "WikiSync"? (a mashup of the two words Wikipedia and synchronous).

After releasing the ad, the wider Wikipedia community added additional ideas (*Tricider Social Ideation and Voting for WikiSync*, 2022).

Social Voting Data Collection and Analysis

The community has voted on ideas and added pros/cons arguments, which others also have voted on, and new ideas and comments in the discussion space.

The Tricider invitation emails were responded to with encouraging replies, bearing witness to the importance of these endeavours. Between the 7th of November and the 15th of December, the survey had over 160 unique views and 11 additional ideas, with 123 votes in total.

Two data sets were used to create the diagrams shown below in Figure 7.32 to Figure 7.34: one from the questionnaire and another from the Tricider tool. The first diagram shows a treemap of nested rectangles representing the ideas. Each rectangle's size and colour intensity correspond to the number of votes an idea has received, ranging from 11 votes at the top left to 1 vote at the bottom right. The ones with no numbers represent ideas added by the community.

Figure 7.32 A treemap of nested rectangles representing the ideas listed in the Tricider.

Since Tricider recently dropped a feature that shows a map of the visitors' countries, the geolocation tag information from the questionnaires was used to map the contribution in the Tricider, as shown in Figure 7.33. The colour gradient refers to the years of experience, while the bottom row refers to the random ID assigned to participants after the questionnaire is done. The following diagram in Figure 7.34 shows the ideas each participant voted on for using the same colour scale in Figure 7.33. Six answers were excluded due to missing data about the Wikipedia experience.

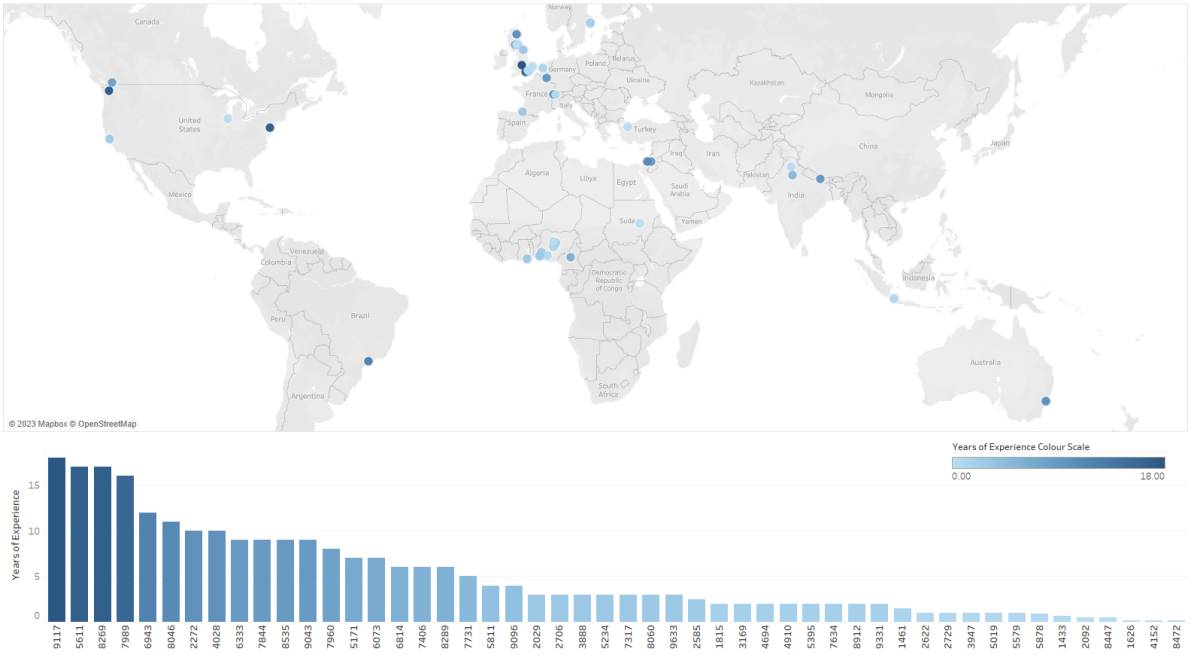


Figure 7.33 Map shows the geographical distribution of the Tricider participants.

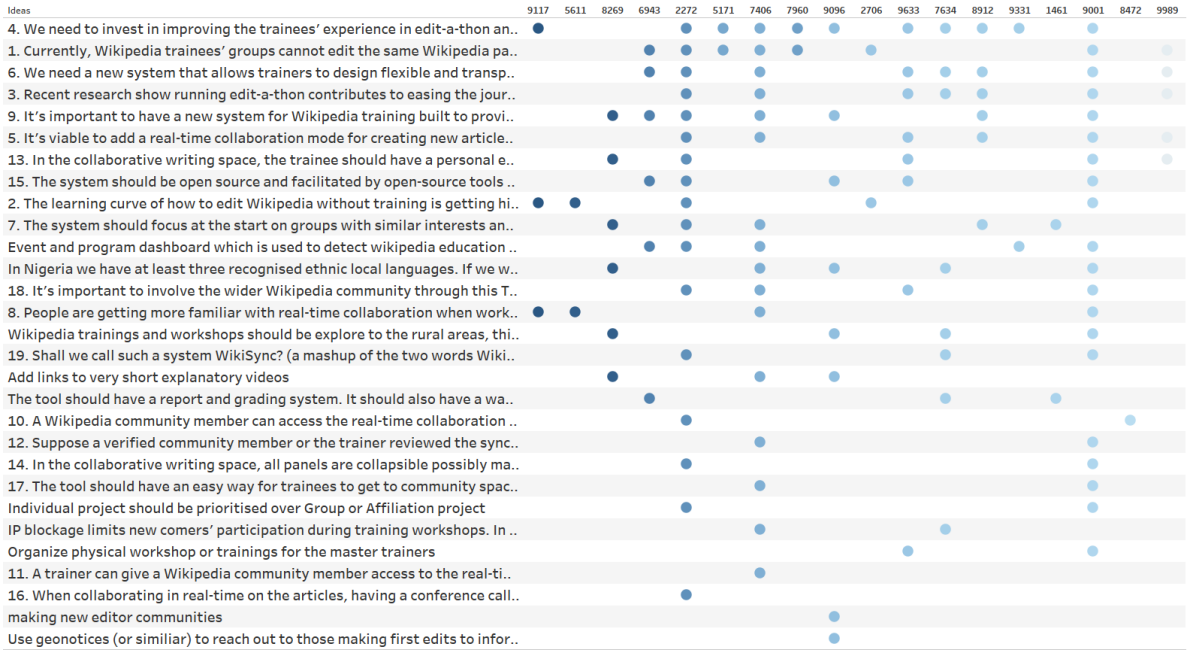


Figure 7.34 The distribution of votes per idea according to the participants' IDs.

The following discussion explores insights from the broader Wikipedia community Tricider's contributions and assesses its correspondence to the sessions' outcome.

The arguments for the initial 19 ideas and the newly added ones echoed most of the discussion points in the sessions, upholding the research direction and output. As seen in the diagrams, the fourth idea about the need to invest in improving the trainees' experience in Edit-a-Thon and Wikipedia training has received the highest number of votes. Ranked second is the inability to edit together, receiving highly voted supporting arguments, such as: "This needs to be resolved. One will get frustrated when you discover that all your edits are gone because somebody has edited the page before you submit yours" P5579.

Idea 6 calls for "a new system that allows trainers to design flexible and transparent wiki workshops to train the small groups on real-time collaborative Wikipedia writing." It was interesting to see this idea, which has the training design interface sketch attached to it (see Figure 7.19), rank third, with eight votes and no arguments. On the other hand, there was an interesting developing debate for idea number two on whether the Visual Editor has made the newcomers' life easier and whether Edit-a-Thons lead to higher retention rates.

Most of the arguments intersected with concerns discussed in the sessions, such as security concerns raised for idea 16 (which calls for access for community members to the conference call to facilitate real-time collaboration).

Similarly, the following newly added idea by the community on Tricider connects back to the IP blocking issue discussed in the Fifth "Create" session, "IP blockage limits newcomers' participation during training workshops. In as much as vandalism prevention is the aim, this should only be applied to those who have received repeated warnings but failed to comply."

Some of the newly added ideas have shaped WikiSync's final prototype changes, as seen in the community interface at the final prototype (see Figure 7.37). For example, the following idea added by P 5811 had a technical requirement asking for "a report and grading system [which] should also have a way recording how many new editors the trainer has been to train and the progress."

Also, several upvoted ideas and supporting arguments mentioned the importance of encouraging minority languages and editing in rural areas.

"WikiSync", as a name for the tool, received three upvotes with a supporting argument.

The last added idea was a call to "making new editor communities, to which newcomers can opt in or out for joining after training to work on topics of shared interests, which could be supported with tools similar to Slack, FB... etc. (or equivalent) and some moderator."

Finally, it is worth mentioning that one cannot judge the importance of the ideas added after advertising the 19 ideas from the research based on votes or arguments. This is because some of the participants who already voted only on the nineteen ideas accessed Tricider before adding them, had no way to vote or discuss newly added ones.

7.3.3 Collaborative Writing Interface High-Fidelity Prototype

Many of the ideas discussed earlier show how the conversation has led to a new design direction for the real-time collaborative space and addressing Q2 What are the key features of a real-time collaborative tool for Wikipedia newcomers? Based on the output of the fifth session, participants' sketching in the iteration session, and Tricider's results, the interface shown in Figure 7.35 has emerged.

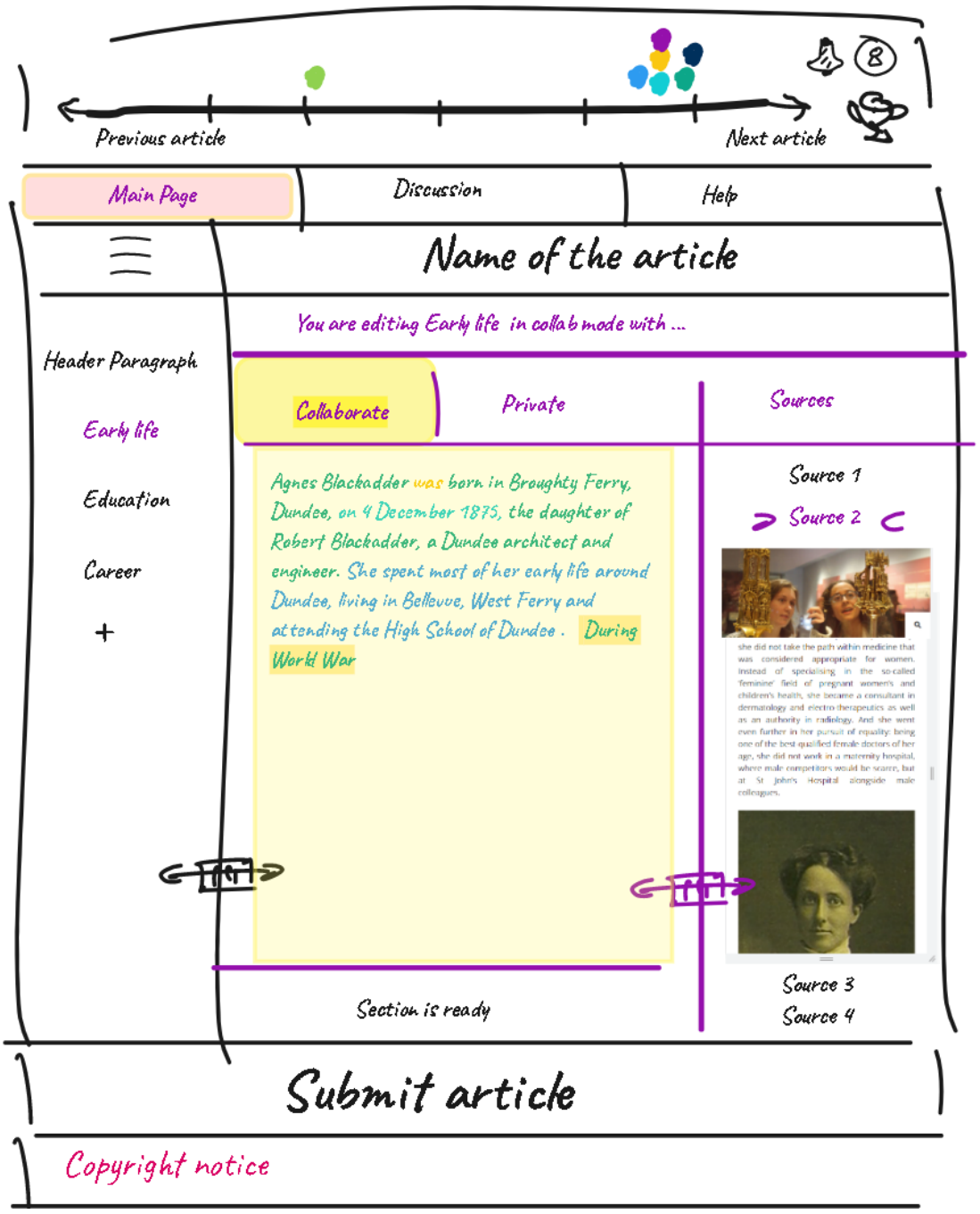


Figure 7.35 Iterated collaborative interface design.

This was then discussed with the research collaborators, leading to the final real-time collaborative writing interface layout concept shown in Figure 7.36, which is divided into three main regions drawn using a solid line. Inside them, the dashed rectangles on the sides are hover-triggered expandable sidebars, whereas the doted blocks serve to illustrate the position of a block of buttons (or actions) for a region.

1. The first **Training Activities Navigation Region** connects the collaborative writing activity to the rest of the training. Through its pagination component, it allows participants to move between articles, be able to identify those working on other articles, the articles statuses (completed or submitted), and whether anyone needs a review or assistance.

   Pressing the help button in this region sends the participants to the main (home) conference call used for the training to discuss any matter with a trainer.

   This region of the interface also allows the participants to revisit earlier stages of the training or refer back to the WikiSync community interface discussed in the conceptual model in Figure 7.12 and the interface shown in Figure 7.25.

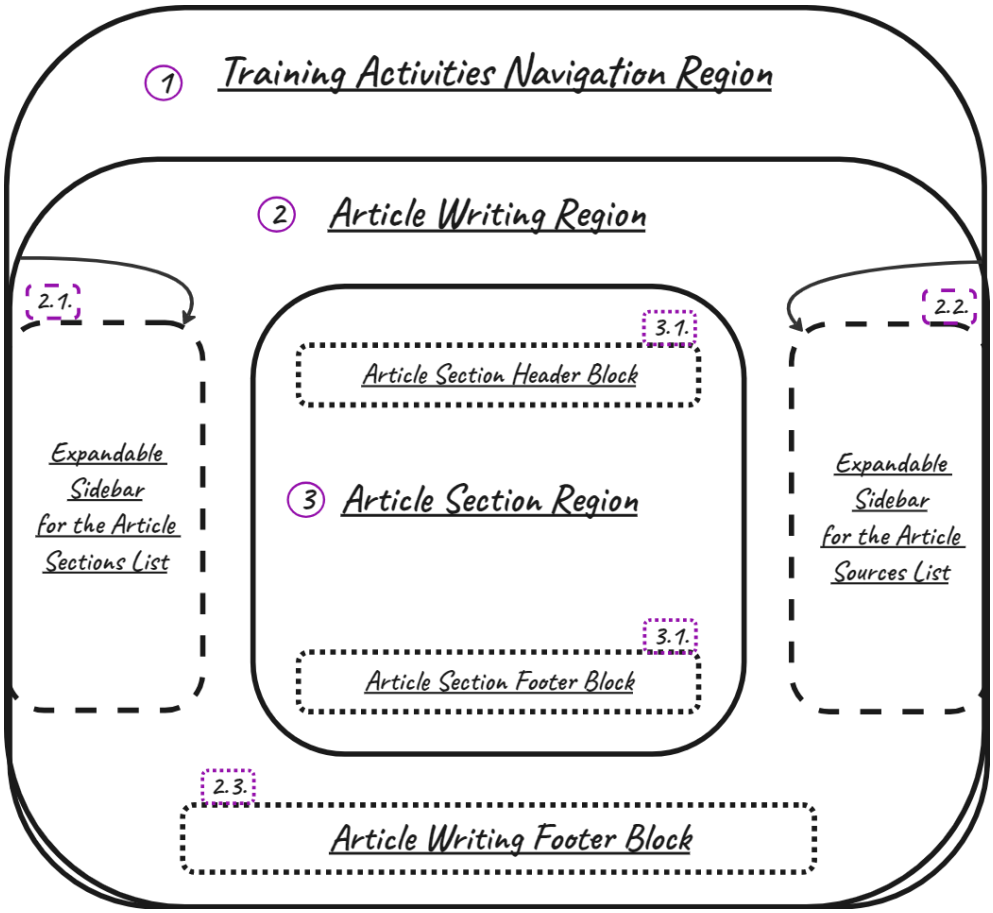


Figure 7.36 The final real-time collaborative writing interface concept.

2. The second **Article Writing Region** allows moving between the Article's Main Page (where the article is being written) and the Article's Talk Page (where the article has an ongoing conference call through an asynchronous chat as a fallback communication channel).

This region has the section writing region in the middle, two sidebars and one actions block at the bottom.

3.1. The left side rectangle is the **Expandable Sidebar for the Article Sections List**, which enables participants to select and move between different article sections, i.e., header, early life, education, etc.

3.2. The right rectangle is the **Expandable Sidebar for the Article Sources List**, which makes it possible for participants to move between sources listed to be used as references for this section or the article as a whole. This sidebar could be pinned and resized to enable participants to avoid having multiple windows open.

3.3. At the bottom, the **Article Writing Footer Block** has a button that could be used to update the article status to be ready for review, submission, or completion.

3. The third **Article Section Region** is intended for writing one section of an article in the real-time collaboration tab or developing parts of the section in separate private tabs. As shown in the figure, this region has two blocks, one at the top and another at the bottom.

3.1. The **Article Section Header Block** at the top could be used to move between tabs (collaborative or private writing). It also has a text styling bar on the left. Clicking the support icon on the right shows a popup interface that has tips or videos that could help with writing the section of the article. The bottom of this popup interface also has a link to the same support icon that leads to the main conference call in case the participants need more help. Finally, a versioning icon next to the support one could be used to roll back to previous autosaved versions of the section's text (depending on its implementation complexity, the last feature might be dropped for WikiSync's first version).

3.2. The bottom **Article Section Footer Block** has buttons to ask the WikiSync community (including a trainer) to review the section, another button to mark the section as complete, and a third one to mark the section as ready for submission. In addition, each section has a call channel, and a speaker icon would be used to mute this audio channel on the bottom right.

To give the participants interaction with the collaborative space, a high-fidelity interactive prototype was built, having two newly-designed interfaces: 1) the WikiSync Community Interface (see Figure 7.37). 2) the Real-time Collaboration Interface (see Figure 7.38).

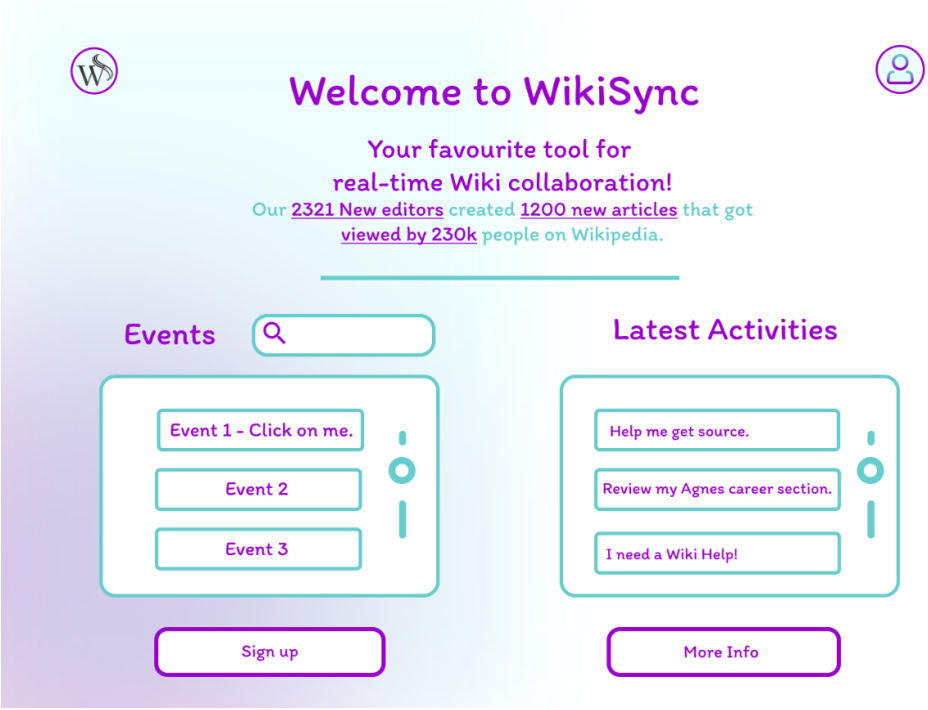


Figure 7.37 High-fidelity WikiSync community space prototype.

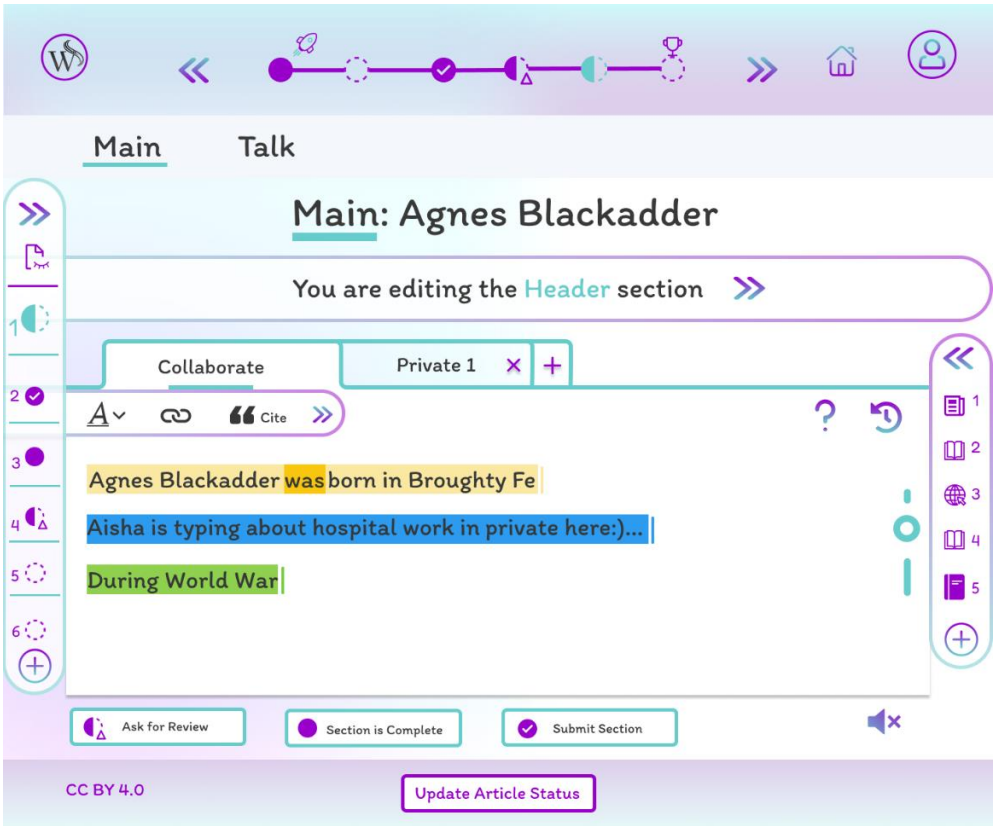


Figure 7.38 An interactive prototype of the last version of WikiSync's real-time collaborative writing interface using Figma.

A link was shared to these interactive interfaces with the participants as part of a Prototype Report (see Appendix F.2), which includes the most important highlights of my research and the process of developing the interface prototype of WikiSync. It was written in a way that helps those who missed sessions to catch up with the rest of the participants. The report included instructions on reviewing the prototype using a live link to the Figma tool, where they could move between interfaces and click some of the interfaces' buttons. For example, once clicked "Event 1" in the WikiSync Community Interface, shown in Figure 7.37, it would send the participant to the Real-time Collaboration Interface, shown in Figure 7.38. Most of the features have been designed to show how some of them would interact with users. Clicking the articles navigation bar, for example, would tell who is working on what article. The left and right expandable sidebars, however, would show the element linked to them, as shown in the design interface of the Figma tool (see Figure 7.39).

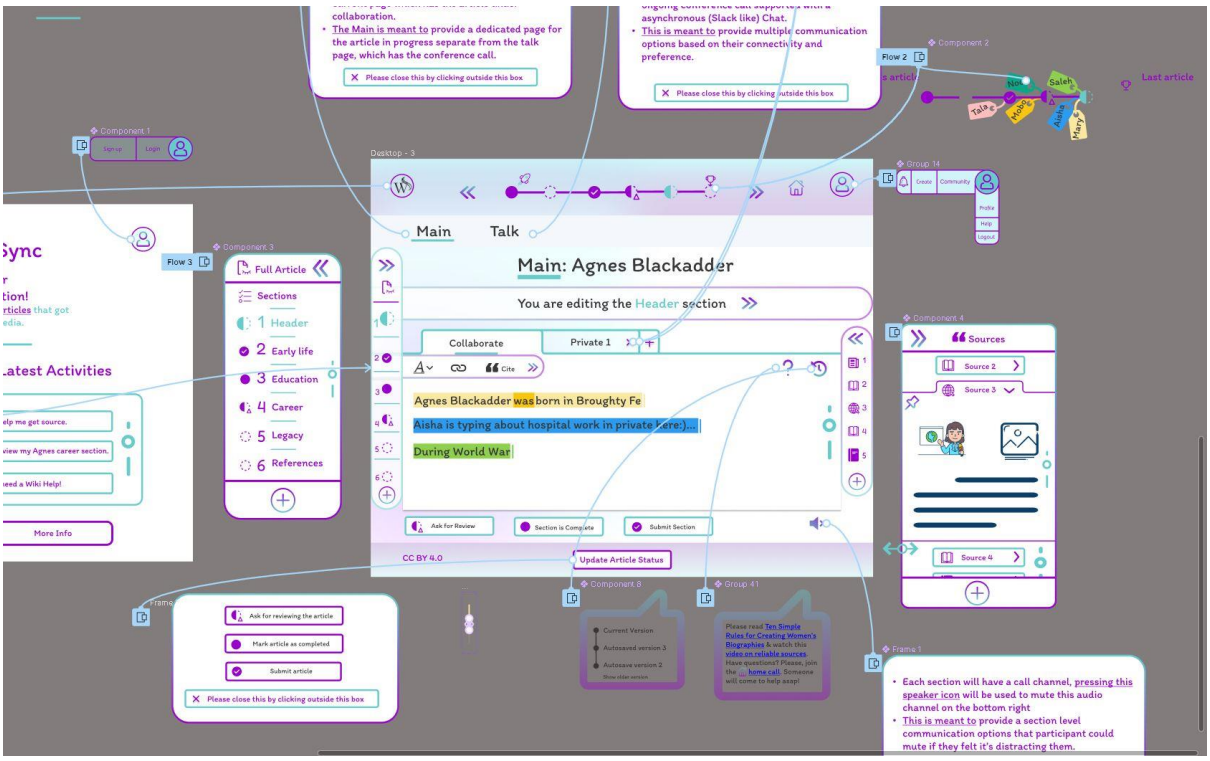


Figure 7.39 Figma project prototyping design space.

The high-fidelity design decisions were based on the data analysis from the sessions. However, in some cases, decisions had to be made based on design heuristics principles, consulting research collaborators and colleagues' feedback, such as the visual appearance, which can be revisited in

future work. In addition, even though most features have been included in the Figma prototype design in the available time window, some features of the Real-time Collaboration Interface were not designed, and clicking on their buttons would only show dialogue windows since these features were not discussed in the sessions.

7.4 Discussion and Summary

The approach followed in collecting data for this research has endeavoured to capture the context of the problem and review the implicit and explicit barriers that face introducing a new model of collaboration. However, even though it was evident that this approach has helped identify several challenges, the more the discussion has evolved, the more complicated the data has become. Thematic analysis, covered in Chapter 6, plays a role in identifying themes that have emerged from these discussions, yet, linking them together has been challenging. This leads to the first question of this chapter.

Q4.1 How to untangle the socio-technical complexity of introducing the new collaborative writing process? Reviewing the literature addressing this question has pointed to the activity theory framework. Building Activity Systems has enabled me to draw a "rich, structured account of work and cooperative work settings" (Turner & Turner, 2001, p. 1). Looking at the broader picture has helped make connections between different components which play different roles in the success or failure of an activity. Therefore, translating part of the data collected during the observational study and sessions into the AS1 newcomers' real-time editing, AS2 onboarding, and AS3 Wikipedia community support activity systems covered in sections 7.1.6 and 7.1.7 has helped untangle the socio-technical complexity of introducing the new collaborative writing process.

Q4.2 How to connect the system requirements with the qualitative results of a thematic analysis? This question emerged quickly after the Conceptualisation phase. The body of literature on IS design, covered in Section 7.1.3, has paved the way for PD practitioners to use activity theory as a tool in their innovation process in fields such as HCI and CSCW. Reviewing the evolution of AT covered in 7.1.1 has been important in identifying how an analysis of contradictions can derive insights and user requirements for WikiSync. The utilisation of the second and third generations of AT has also been vital in defining the different levels of contradictions for driving system requirements and insights out of the thematic analysis and building the network of activities covered in Section 7.1.8.

Q4.3 How and when to scale up and down participation in designing a tool that serves a large and widely distributed community?

This question has been hard to answer. Botero et al. (2020, p. 17) discuss how going through user participation in the PD process is "enacted in a broadening array of circumstances" and requires skills beyond the ability to select the right methods for PD projects to include aspects such as being aware of the strategic and mundane issues that emerge during the PD process. In my research, several factors have played a role in deciding when to iterate, consult with the wider community, when to refer back to the literature for guidance or answers, and how far to engage the PD participants in the decision process. The following are some of these factors that should be considered in **scaling-up and down participation** in designing:

1. **Time**. this is a two-dimensional factor:
    a. Managing the relationship with my research collaborators (Wikimedia UK and others), in addition to the participants, has been challenging. Full participation in the design process requires an insurmountable amount of time. The availability of participants was irregular, being dictated by scheduling conflicts. For example, some were available during the week, and some on weekends. Therefore, trade-offs and prioritisation had to be made.
    b. A different way to view ***time*** as a factor is to assess how much time it can take to explore, adapt, and deploy the methods or analytical tools. As discussed in Section 7.1, pragmatism had to be used in managing the interplay between searching for methods and preserving the research's momentum.
2. **Technological Considerations.** Including factors such as:
    a. Participants' technological capabilities. Choosing the tool for involving the wider community has been an important factor in the success of the process. Using a Wikipedia page could have limited the discussion to the community who monitor innovation discussion spaces and Wikipedia tools, such as Wikieditor. Thus, Tricider was chosen as a middle ground. It has both the democratic participation that the community respects and is easy to navigate at the same time.
    b. Participants' access to devices that enable the co-design. Some participants were able to join only through phones or tablets, even though they were asked in the invitation

emails to get access through a PC, to achieve maximum benefits from their experiences when working on Miro boards.

c. The reliability of the tools used in the innovation process. For example, between the time of choosing Tricider, as part of the innovation process and that of using it, three features malfunctioned.

d. The license and data processing practices of the tools used in the innovation process should be carefully considered. Users, in general, are becoming more vigilant about how their data are being processed, and the Wikipedia community has paid much attention to this issue, as well as to who has the right to use it. This has led to giving proper weight to how Tricider treated data and flushing it out for participants to clear any confusion with regard to their participation in this tool.

3. **Feature's Conceptual Model Maturity.** Guided by the PD principles covered in Chapter 2, this research has centred around including users in shaping solutions that fit their social and technical environment. The aim is to co-design a highly accepted and sustainable approach to introduce real-time collaboration and possibly influence the innovation process of the broader community.

   However, the question to be asked is when the broader community should be involved and how much the participants had to be iterated on the design. The focus had to be on the common understanding of ***solution maturity*** among my research session participants. I argue that before involving the broader community at an early stage in long debates, or an "unwarranted [community member] bias" (Hartson & Pyla, 2012, p. 146) whether a feature could be implemented or not, one had to ensure addressing the most critical concerns of the community members involved in the sessions. Such extensively long debates can be seen in the 15 archives of Wikimedia's Flow (or Structure Discussions) feature (*Flow Project Talk Page - Archive Index*, 2023). The Flow project, discussed in Chapter 2, was criticised by the community, leading to its removal from the English Wikipedia MediaWiki, in 2016, upon a request from the community with "no plans for it to return" (*Wikipedia Flow Project Main Page*, 2022), leading to a new consultation process with the community, in 2019 (*Talk Pages Consultation 2019 - MediaWiki*, 2019).

   The success in introducing change and eliciting computer specification makes it necessary "to do more than 'just add users and stir'." (Muller & Druin, 2002, p. 3).

Additionally, I would argue that community involvement early in the PD process should not be limited to specific projects that the designer thinks need community involvement. Changing any feature that affects the community norms should be consulted starting from the Discovery phase in a small inclusive group, leading to the stage where the wider community is consulted. To an extent, Wikimedia has already done this in their new community consultation and recent strategy (*Movement Strategy Recommendations*, 2022). However, the remaining challenge is making such a process systematic in such a way that the debates do not spiral out of control and overwhelm the organisation's staff members (designers, strategists, or IT specialists). A single Wikipedia debate can span "several years and hundreds of thousands of words" (Jemielniak & Przegalinska, 2020, p. 52). In addition, such debates and the instinct to win an online discussion in Wikipedia are associated with creating a "conflict-driven environment, and a clear gender bias among many editors result in a huge gender gap", as well as "burnout among peer production project members" (Jemielniak & Przegalinska, 2020, p. 51 52). Furthermore, such debates may allow experienced editors to use Wikipedia's rules to "re-frame the debate in terms that favour some participants over others, to regulate available and acceptable discussion topics, and block participation by editors that hold minority views." (Morgan et al., 2012, p. 3497).

My research is not a call to move designing tasks into "back office" responsibilities, limiting the size of user involvement in the design, or questioning the "wisdom of the crowd". Instead, I argue that assessing and finding the right phase and size of the ***crowd*** in which their participation in the innovation process could drive it forward rather than cease its progress or direct it for the benefit of one group over the other. Therefore, in some cases, driving the broad community into such debates over the platform changes in an early phase can defeat the purpose of involving a diverse representation of the community.

Prior to agreeing on the need to involve the broader community, the participants and I discussed major issues regarding the introduction of real-time collaboration that could have easily turned into dull, lengthy debates. Looking back at the Social Voting phase's arguments, votes, as well as new ideas discussed in Section 7.3.2, it was interesting to see some of the ideas presumed to attract the long debates did not provoke any arguments. For example, idea 6, which calls for "a new system that allows trainers to design flexible and transparent wiki workshops to train the small groups on real-time collaborative Wikipedia writing.", ranked third, with eight votes and no arguments. This contrasts with any existing innovation processes covered in Chapter 4, in which

the community discussed real-time collaboration. Of course, other factors could have played a part in reducing the debates, such as the use of the Tricider tool instead of a Wikipedia page, the size of the audience attracted through the ads, and the fact that real-time collaboration would be external to Wikipedia. However, this also demonstrates that the **maturity of the feature's conceptual model has a significant influence on assessing the scale up/down participation in designing and should be evaluated in collaboration with the PD participants.**

This chapter has covered the last phases of the PD process (The Design and Social Voting phases). First, the design concepts were analysed with the help of AT as a means to develop the first sketches of WikiSync, leading to the iterative process involving the wider community in shaping the final prototypes. Activity Theory has helped me work out how different activities interplay in a "what-if" scenario of a training session that involves real-time collaborative writing. This led to a culturally aware discussion of the challenges in the contradiction shown in Table 7.1 and Table 7.3 and producing design requirements for WikiSync. The high-fidelity prototype, shown in Section 7.3.3, introduces a real-time collaboration feature to Wikipedia training and article writing, which has gained community acceptance and opened up the future discussion of implementation, which is covered in Chapter 8.

Chapter 8 Discussion of Contributions

My research started by exploring past design decisions that constrain the current collaborative writing model options in Wikipedia training. Although it is technically possible to add real-time collaboration, it has not been possible to implement it due to barriers arising from Wikipedia's long-established traditions and the commitment to the liquid collaboration model. Despite the usefulness of real-time collaboration, trainers in synchronous training sessions were observed asking newcomers to adopt an asynchronous collaboration style – as if they were not co-located in time. This shows one of the many ways Wikipedia is locked into its existing collaboration style. The path dependency of the wiki-based editing model makes changing the socio-technical system a real challenge.

As Ananian et al. (2018) repeatedly highlighted at three different Wikimania conference presentations between 2015 and 2018: "Lots of talk among developers about technical mechanisms for real-time collaboration. Very little talk about the social aspects or impact on community. Not much dialog between developers and editors and readers. Let's start to fix that!". Thus, my PhD research aims to find ***how*** to involve Wikipedia's community and newcomers in designing a solution for introducing real-time collaboration while being responsible and respectful of Wikipedia's rich social structure and history. A prototype of the solution emerged through the pursuit of addressing this question.

WikiSync is a prototype for a system that supports synchronous collaboration in Wikipedia training and editing events. It is a transparent, flexible environment for real-time collaborative writing that can be used to support mediated contributions to Wikipedia. The conceptual model for WikiSync presented in Chapter 7 addresses critical social challenges that have emerged from the observational study and discussions covered in Chapters 4 to 7, such as authorship, transparency, reverting content, reviewing process, or licensing content.

WikiSync targets the Wikipedia training environment. A more general solution that introduces real-time editing to all of Wikipedia would at least require more research on Wikipedia's innovation process and community norms. The distributed PD framework used in this research can be used to study and evaluate the influence of liquid collaboration on power structures and on content as well as how real-time and asynchronous writing could co-exist. Findings from

the study are important for developing a possible comprehensive solution for introducing real-time collaboration.

Therefore, this research makes two major contributions: WikiSync, a co-designed prototype for a system that supports synchronous collaboration in Wikipedia training, and a distributed co-design framework which can be used in distributed diverse communities, similar to that of Wikipedia.

My research contributions are discussed in detail below:

Contributions C1 and C2 are based on what is covered in Chapters 2 and 4 and connect to Q1.

C1: Exploration of the potential of introducing new real-time writing technology in Wikipedia training events.

A synthesis of related research work is provided in Chapter 2, discussing the benefits of real-time collaborative editing and existing guidelines for its tools outside the Wikipedia context. The literature review covers different forms of collaboration. Since I could not find research on introducing real-time collaboration to Wikipedia, I reviewed the literature on using Wikis in other contexts, such as education, to learn about previous attempts to introduce real-time editing. This showed that, while tools such as real-time chat were added to Wikis, real-time editing functionality was not included.

Chapters 2 and 4 demonstrate a gap in research on introducing new collaborative technology in Wikipedia training events and its effect on group dynamics, specifically focusing on trust and its influence on the writing approach.

The co-design project making up the bulk of this research reveals a need to research the evolution of collaboration models, such as the changing role of the Talk Pages. A further topic for research yet to be done is the interplay between Wikipedia and other technologies, such as Facebook and WhatsApp, which editors often use to coordinate their activities. Future research should look beyond Wikipedia to study the wider ecosystem and its influence on collaboration models.

C2: Analysis of research findings on the evolving *liquid collaboration* model in Wikipedia and its links to bias in the content.

The approach followed in collecting data ensures capturing the context of the problem while reviewing the barriers hindering the creation of a new model of collaboration. Additionally, it

followed Hagen et al.'s (2012, p. 7) recommendation for shaping the PD process through end-users' direct involvement and insights from existing scholarly work. This has led to exploring several challenges that Wikipedia faces. Specifically, scholars have discussed collaboration constraints among newcomers in Wikipedia's liquid collaboration model described by Jemielniak & Raburski (2014). This is increasingly a problem for new editor retention (Halfaker et al., 2013) and contributes to the prevalence of low levels of trust (Jemielniak & Przegalinska, 2020) and antagonistic behaviour, which contributes to continuing bias (Collier & Bear, 2012). Researchers such as Shane-Simpson & Gillespie-Lynch (2017) and Langrock & González-Bailón (2022) have shown how the potential of new models of collaboration, feedback, and Wikipedia training may address bias.

My work provides evidence that members of the Wikipedia community agree with this assessment. This is visible not only in the co-design sessions, but also through the Social Voting and Ideation phase covered in Chapter 7. For example, the response from the wider community was supportive of introducing real-time collaboration into Wikipedia training through WikiSync, which has been co-designed throughout the research. WikiSync allows for transparency and flexibility in getting Wikipedia community members to access and support the training. Such an approach was discussed as a possible solution to the ongoing problem of reverting newcomers' work and even getting their IP blocked, especially in a range of IP addresses block[123] which could affect up to millions of users. Such a problem was more apparent with Nigerian participants.

The literature review in Chapter 2 forms a picture of why real-time collaboration is missing. Later in Chapter 4, the investigation of Wikimedia's innovation process complemented this picture by mapping the challenges of innovation in Wikipedia, focusing on previous community attempts to introduce the feature. Each proposal has a unique view of why real-time is needed in Wikipedia and how it could be implemented, with members diving into the technical details of real-time collaboration features. However, as Ananian et al. (2018) highlight, there was little evidence of work on the social impact of this feature.

Chapter 4 demonstrates how the current collaborative model is a barrier to Wikipedia trainees and trainers who are in synchronous training sessions but must work asynchronously. Furthermore, there was confusion in the switch between Wikipedia's Source and Visual editing

[123] https://en.wikipedia.org/wiki/Wikipedia:Blocking_IP_addresses#Range_blocks

tools. Participants even tried to edit in real-time through screen sharing, but in some cases, trainees' lack of technical know-how impeded the collaboration.

Contributions C3, C4 and C5 are based on what has been covered in Chapters 5 to 7 and connect to Q2, Q3, and Q4.

C3:Providing a prototype of a possible solution to introduce real-time collaboration through Wikipedia training.

In my thesis, I provide a prototype for a tool called WikiSync, the first Wikipedia training tool that includes real-time collaborative editing of Wikipedia articles. Co-designed using a distributed approach, it involves the Wikipedia community through several phases that vary in focus and scope of user participation.

Having insights from an observational study covering community views on the current collaboration model and alternatives have helped restart an informed discussion about real-time collaboration. The activities in the sessions covered in Chapters 5 to 6 led to the collaboration model and prototype described in Chapter 7. The WikiSync model addresses community concerns about real-time collaboration and delivers functionality addressing additional needs, such as flexibility in the training design.

By co-designing an external tool that functions as a mediator between Wikipedia and the training participants (newcomers and trainers), WikiSync allows for transparency and flexibility in getting Wikipedia Community members to support the training. The process of running a training using WikiSync (discussed in Chapter 7) highlights how the Wikipedia community can play a significant role in the WikiSync community by supporting its trainees.

The real-time collaborative scenario is delivered in an easy-to-deliver format that resonates with the participants' experience. It addresses the challenges discussed in the sessions, leading to a workable solution in the specific context of Wikipedia training. The acceptance of real-time collaboration in training may open up the possibility of adopting this collaboration model elsewhere, such as in WikiProject.

C4: An adaptive framework for co-design solutions for small research groups in a distributed setting.

While the project was focused on a specific design intervention in the Wikipedia community, as a result, I provide in this thesis a generic framework for co-design solutions for small research groups in a distributed setting. Using an observational study and an adapted version of Hagen et al.'s

(2012) PD framework, I have developed an adaptive, carefully structured co-design approach customised for involving a widely distributed online team. Engaging the participants over time has helped capture a wide range of community concerns, resulting in a mutual learning environment that tackled a complex socio-technical question. After fostering an in-depth community discussion, the research was scaled up to consult the broader community using an online social debate and voting tool.

Drawing on work in Computer-Supported Collaborative Work (CSCW) and Participatory Design, this thesis provides a set of tools and practices for such co-design for a wider community. The multiscale framework developed shows how a clear ethnographic understanding of the research context, covered in Chapter 4, enables both the co-design in Chapters 5 and 6 and the subsequent engagement of the wider community, outlined in Chapter 7.

The resulting distributed co-design framework can be used even by small teams with limited resources to design tools at the scale of a worldwide community by using different methods appropriate for each phase of the project. This is crucial to meet the challenge of introducing socio-technical change in today's platforms, especially a widely distributed, community-governed platform such as Wikipedia.

Hagen et al.'s (2012, p. 1) framework "has been developed with this flexibility in mind, [since the authors] expected that the methods and approaches proposed in [their] guide may be combined with other methodologies". Therefore, instead of reinventing co-design principles and best practices, their framework provides a flexible, phasic structure for my research to build on. My framework, however, differs from Hagen et al.'s in eight key respects:

1) It starts with an ethnographic study of a community and its practices. This was deemed necessary to establish familiarity with the community rules and practices, as these are a crucial part of how the community operates. Of specific interest was training, as it serves as an introduction to the community for new editors.
2) Similarly, a new phase is added at the end to scale up and consult the wider community. This ensures the inclusion of a broader range of voices and acceptability in the wider community.
3) Hagen et al. do not specify how or when to iterate a phase. Because of the resource constraints in this research, the decisions are made based on the state of co-design. For example, as discussed in Chapter 6, there was a need to review the research direction after

the “Position” session, leading to a more inclusive PD process. Also, when the material designs of the interfaces started developing, the need emerged for multiple iterations to address community concerns.

4) My description of the PD process reveals in greater detail the mundane work of running a PD project especially online. Furthermore, it highlights adapting methods dynamically to the unfolding needs of the project.
5) There is a need to describe the analytical tools’ output (thematic analysis and Activity Theory) in an easy format to communicate to participants. In my research, for instance, the “what-if” scenario for the collaborative persona is converted into a storyboard, gaining participants’ acceptance.
6) The challenge being in designing a complex socio-technical system, there is a need for a higher level of transparency and accountability, without which the designed solution does not find acceptance in the wider community.
7) As discussed in Chapter 7, part of the user-designer relationship should involve evaluating when to expose the design to the wider market or community. My research demonstrates that the maturity of the solution conceptual model affects assessing the scale up/down participation in the PD process and should be evaluated in collaboration with the PD participants.
8) User participation does not automatically lead to acceptance of new designs, as a crucial component is building a strong relationship between all the participants (Heiskanen et al., 2010). Therefore, my approach involves building a safe space for participants, ensuring inclusivity as well as caring (and sometimes advocating) for the participants, such as supporting individuals to make and follow the progress on their contributions.

C5: The application of the distributed co-design framework to Wikipedia’s innovation process.

My final contribution is to demonstrate the value of the framework developed for the Wikipedia community specifically. This research has shown promise to provide new impetus for innovation in Wikipedia. For example, as discussed in Section 7.4, the scale up/down participation in designing new features can help Wikimedia ensure inclusivity and transparency by opening it to everyone while strengthening it against unsustainable debates. The focus on the design process

and on tracing concrete decisions to contributions by participants aligns with Wikipedia's traditional open decision-making practices.

Adopting my framework would help lift the level of community engagement in the design process from informing and consultation to higher levels of partnership and participatory design. The framework has a focus on inclusiveness, and it is therefore an appropriate response to the kind of challenges Wikipedia is facing.

8.1 The Generalisability and Use of the Framework

Section 3.4 covered the framework process and its phasic approach. This section is a follow-up, providing a scope of use and structured representation of the methods and tools used to drive the participatory design activities and analysis. Figure 8.1 illustrates the EDPD framework components while combining the methods and techniques used throughout this thesis activities. Activities are outlined approximately in parallel to the phase in which it was used, yet loosely to preserve flexibility and invite other researchers or professionals to adapt it to their needs. The letter (G) next to the activity signifies that it is recommended to divide the working participants into subgroups. The letter (A), on the other hand, indicates the activity's suitability for an asynchronous form of participation. As shown, iteration of the phase should take place only in case of unaddressed concerns that increase in velocity in the last three.

The smaller the team, the more helpful it would be to have an expert or advisory panel with the necessary expertise to consult on the process, starting from the early stage of defining the initial questions to implementation. Their participation would vary based on such factors as expertise, availability, and how critical the situation at hand is. Such consultation could be valuable in designing the workshops and activities or reviewing the results of the analysis stage. However, involving them in all details could slow the project and may result in a loss of momentum. Therefore, it is highly recommended that they have a balanced involvement with the minimum of helping the facilitator with the decision to move from one phase to another.

Once participants finish the session, it is important to keep an asynchronous channel open, such as email communication or a space dedicated for feedback on the session's Miro board, to get their feedback on the session activities and any produced results, in addition to the decision to move from one phase to another.

The diagram shows an additional PD Phase 5 Use that covers the use and evaluation of an implemented system. While this is not covered in this thesis, it is a logical extension of the phasic

model. Activities and methods that can be utilised in this phase are shown similarly to other phases. The dotted arrows in PD Phase 5 Use denote that receiving feedback and experimenting with the tool may result in revisiting Phase 3 Design and iterating with the participants on new versions to be tested in Phase 5. If the changes are significant, this may need yet another consultation with the broader community in PD Phase 4 before moving to Phase 5 again. Finally, Appendix G provides a guide of several steps on how to run a project with a similar context to this PhD research.

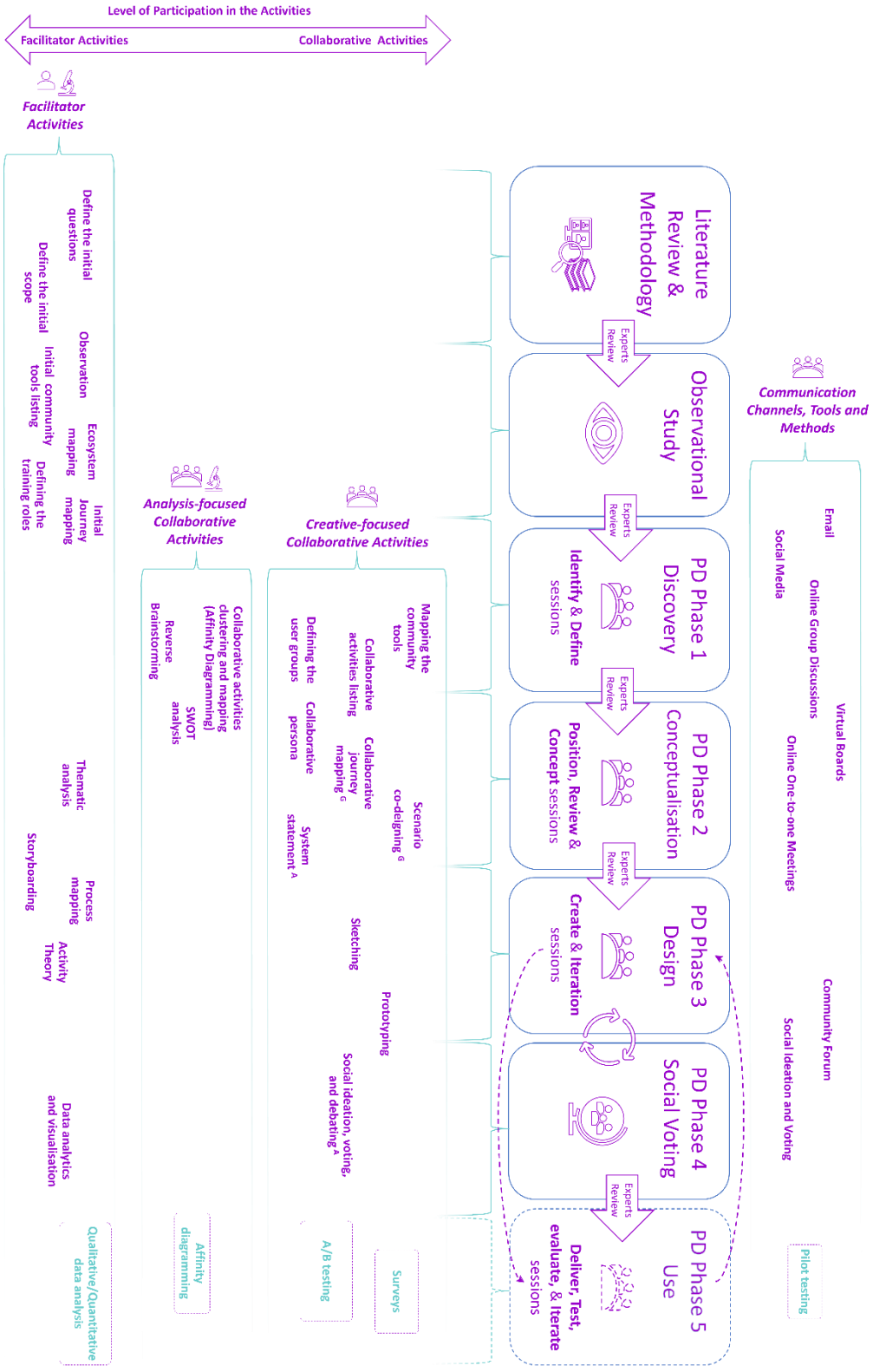


Figure 8.1 Ethnographically-informed Distributed Participatory Design framework.

The proposed framework has been designed to tackle significant change in a complex socio-technical system, although the resulting complexity of the framework is unsuitable in the small one.

Below are four contexts. Applying the framework in the first context has excellent potential in producing solutions with high acceptance and success, while the second, third, and fourth contexts require further adaptation to fit the framework and paying attention to the politics in introducing the change.

1) Non-profit platforms run in partnership with a community of volunteers. Applying this framework can be valuable to the OpenStreetMap (OSM) community,[124] Translators without Borders[125] and the Khan Academy Support Community.[126] Furthermore, community-run software projects that aim to establish platforms, such as Mastodon[127] or PeerTube[128] face similar issues, as they must adapt to increased uptake by people whose expectations are shaped by corporate platforms.
2) For-profit platforms run in partnership with a community of volunteers: This applies to community-led for-profit platforms, such as Reddit and Stack Overflow. However, it is contingent on having an empowered community and corporations that are supportive of their communities.
3) Brand communities: This framework supports facilitators considering a new user-inclusive innovation approach in the business and for-profit environment. While it can be used to explore solutions with these corporate communities in the brand periphery, such as the LEGO Ideas community[129] or Apple Support Community,[130] using it requires careful consideration of the power dynamics involved.
4) People-centric organisations: This framework can benefit organisations seeking a decentralised decision-making environment and building an employee-community-driven culture. It aids User Researchers and Product Managers' understanding of context better, ahead of bringing cross-functional teams together with users in developing their products.

[124] https://community.openstreetmap.org/
[125] https://translatorswithoutborders.org/
[126] https://support.khanacademy.org/hc/en-us/community/topics
[127] https://joinmastodon.org/
[128] https://joinpeertube.org/
[129] https://ideas.lego.com/community
[130] https://discussions.apple.com/welcome

Similar to the second and third contexts, this one also demands paying great attention to power dynamics and process transparency, as these are key elements in building the trust needed to get participants' ownership, as successful solutions are devised.

8.2 Limitations and Open Questions

8.2.1 Representativeness

My research focuses on experienced trainers and editors with limited participation by newly trained Wikipedia editors. Recruiting newly trained Wikipedians has proved difficult, and I had to cancel some sessions due to the participants' absence. Even though careful analysis and expert reviews were involved in the design decisions, the limited involvement of the newly trained Wikipedians might have resulted in bias in the conceptualisation of the tool, shifting the direction in favour of the trainers or experienced editors. The research could have also benefited from the participation of non-users or those who resist using Wikipedia. The PD social voting, open to the global community, was presumed to attract non-users. However, the responses show little evidence of their participation.

Given the voluntary participation, those who have participated in the sessions or social voting and ideation may have strong feelings towards introducing a new collaboration model. However, this also may have attracted those who are sceptical and resistant to this change and the efforts that challenge the status quo. The group of participants, therefore, have been more partisan on key design issues than the community as a whole, nevertheless, an agreement has been reached on a design that addresses a broad range of concerns.

8.2.2 Open Design Issues

Further research on the following elements is needed to improve the WikiSync design:

- WikiSync focuses on creating new articles and does not deal with editing existing ones. Future iterations or new "what-if" scenarios could be expanded to address editing existing articles in real-time, along with asynchronous editing.
- The third iteration of the "Create" session in the Design phase discusses the success of the design in considering all the concerns discussed previously. However, Participant B mentioned the need to advance abuse prevention measures in WikiSync, such as coordinating such efforts with the Wikipedia community and benefiting from their user blocklists. This is

discussed partially in this research session. Future work to explore it further can potentially improve the WikiSync system concept.

- The need exists to clearly define a list of rules on topics, such as: what would happen when an article is marked “complete” or submitted for review, who has access, and who suggests edits in WikiSync and Wikipedia communities.

8.2.3 Future Work

Most research on PD projects does not move beyond design and into implementation (Korsgaard et al., 2016). In the same vein, WikiSync is not implemented as the complexity of intervention in a complex socio-technical system such as Wikipedia puts implementation well beyond the scope of a PhD research project. Nonetheless, significant effort has gone into the co-design of a prototype and into ensuring that the features it represents are grounded in a thorough understanding of the Wikipedia community, its norms and practices. They have been developed through a PD process that balances the need to have depth and breadth in participation.

This thesis illumines the problems that the implementation of real-time collaboration must solve; indeed, my research provides candidate solutions to several of those. An implementation effort would need to continue in the same spirit as the PD process initiated in this research to ensure that the community is involved in design decisions. While the system would be separate from Wikipedia, agreements on the use of real-time editing functionality would need to be established. For example, an agreement must be reached on real-time articles being produced through WikiSync in terms of reviewing and author attribution.

Future work on rollout into the community could look into the following:

- Discussing the use of WikiSync in a co-located vs online and co-located then online.
- Exploring the training sessions mediating artefacts that will be used through WikiSync, such as the conferencing system.
- Studying how the system can shape and foster the WikiSync community, as well as studying what feature affects the level of trust community. This connects to this proposed idea from the PD Social Voting Phase (*Tricider Social Ideation and Voting for WikiSync*, 2022):

  > "**Making new editor communities:** Attendees at wiki-editathons and newbie training have common interests, there should be an easy way to make community groups spring up from the event (opt-in or opt-out) or allow peer

support after the training. This may require other tools like Slack, FB etc (or equivalent) and some moderator."

My research lays the groundwork for future research to investigate Wikipedia training group dynamics and the impact of technological choices on issues such as systematic bias. Moreover, the relationship between the innovation process and the collaboration model and their effect on content bias should be investigated further. Also needed is advancing our understanding of Wikipedia's mix of innovation paths and how their community interacts with different possibilities for involvement and asking whose voice is promoted and whose is silenced as a result of using a certain mix of user innovation paths. This includes considering involving the non-users and readers of Wikipedia and studying the influence of innovation trends in the wider eco-system on Wikipedia.

Making design decisions that work for a wide range of users is a challenging task. Additionally, the very large userbase makes it conceivable that quantitative methods might be used to generate decisions based on feedback from users or through design experiments, as in Dudley et al.'s (2019) research using the Amazon Mechanical Turk to improve interface design. There is a space between participatory design, user experience design, user innovation and quantitative methods, such as data-driven user research. Innovation in this space may be needed to bring broader changes to the Wikipedia ecosystem.

8.2.4 Concluding Remarks

Collaboration on Wikipedia is limited by Wiki technology, allowing asynchronous collaboration only, while synchronous editing is discouraged by the Wikipedia community, as it can result in excessive merging conflicts and other problems. Attempts to add real-time collaboration have faced difficulties which are deeply rooted in Wikipedia's culture, community norms, and innovation process.

The aim of this research is to advance an understanding of involving Wikipedia's community and newcomers in designing a solution for introducing real-time collaboration while being responsible and respectful of Wikipedia's rich social structure and history.

My research has analysed Wikipedia's collaborative writing models and the potential for change, especially in training events, helping co-design key features of a real-time collaborative editing tool for newcomers with a detailed scenario, proposing the first step for fostering real-time

article writing through an external system, WikiSync. Further collaboration with the community on implementation can ensure scale and sustain the solution and open space for new research, such as studying the collaboration model effect on content bias.

Based on Hagen et al.'s (2012) framework, this research puts forward a flexible framework with online methods and activities, helping small research groups co-design solutions in a distributed setting. It provides a detailed account of the mundane work of doing PD for online communities, including a demonstration of the necessary preparation work in the sessions and the ways to provide an environment of mutual learning between the designer/s and participants. Finally, this thesis offers a detailed methodological stance as a foundation to help distributed community researchers devise their approach to study and introduce change in complex socio-technical environments.

Bibliography

T196211 Deploy CollabPad to officewiki. (2022). https://web.archive.org/web/20220927193535/https://phabricator.wikimedia.org/T196211

T241180 RFC: Adopt a modern JavaScript framework for use with MediaWiki. (2022). https://web.archive.org/web/20230207222827/https://phabricator.wikimedia.org/T241180

Abdullah, Z. (2014). Activity Theory as Analytical Tool: A Case Study of Developing Student Teachers' Creativity in Design. *Procedia - Social and Behavioral Sciences*, *131*, 70–84. https://doi.org/10.1016/j.sbspro.2014.04.082

Agnes Blackadder Hall—Wikipedia. (2018). Wikipedia. https://web.archive.org/web/20230304220328/https://en.wikipedia.org/w/index.php?title=Agnes_Blackadder_Hall&oldid=871910026

Agnes Forbes Blackadder. (2022). In *Wikipedia*. https://en.wikipedia.org/w/index.php?title=Agnes_Forbes_Blackadder&oldid=1101123438

Al-Ali, S. (2020). Activity systems analysis: A maze worth exploring. *Studies in Technology Enhanced Learning*, *1*(1). https://doi.org/10.21428/8c225f6e.39ed7528

Alexa - Top sites. (2022, March 3). *Alexa—Top sites*. Alexa. https://web.archive.org/web/20220303060420/https://www.alexa.com/topsites

Allen, C. D., Ballman, D., Begg, V., Miller-Jacobs, H. H., Muller, M., Nielsen, J., & Spool, J. (1993). User involvement in the design process: Why, when & how? *Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems*, 251–254. https://doi.org/10.1145/169059.169203

Ananian, S., Chan, T., Kattouw, R., & Sanders, E. (2018). *Program/Real-Time Collaboration: Connecting Humanity—Wikimania*. https://web.archive.org/web/20221202172130/https://wikimania2018.wikimedia.org/wiki/Program/Real-Time_Collaboration:_Connecting_Humanity

Anderka, M., Stein, B., & Lipka, N. (2012). Predicting quality flaws in user-generated content: The case of wikipedia. *Proceedings of the 35th International ACM SIGIR Conference on*

Research and Development in Information Retrieval, 981–990. https://doi.org/10.1145/2348283.2348413

Andreini, G., & Forbes, P. (2021). *TRIPLE Deliverable: D3.4 Report on Co-Design of User Profiles and Dashboard Research*. https://doi.org/10.5281/zenodo.5702098

Androutsellis-Theotokis, S., Spinellis, D., Kechagia, M., & Gousios, G. (2010). Open Source Software: A Survey from 10,000 Feet. *Foundations and Trends in Technology, Information and Operations Management*, *4*. https://doi.org/10.1561/0200000026

Arabic Wikipedia's Stub Articles. (2023). Wikipedia. https://web.archive.org/web/20230216003216/https://ar.wikipedia.org/wiki/%D8%AA%D8%B5%D9%86%D9%8A%D9%81:%D8%AC%D9%85%D9%8A%D8%B9_%D9%85%D9%82%D8%A7%D9%84%D8%A7%D8%AA_%D8%A7%D9%84%D8%A8%D8%B0%D9%88%D8%B1

Arazy, O., & Nov, O. (2010). Determinants of Wikipedia Quality: The Roles of Global and Local Contribution Inequality. *Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work*, 233–236. https://doi.org/10.1145/1718918.1718963

Ardati, A. A. (2018). *System Analysis for an Online Volunteering Collaborative Translation Platform that Supports Translating Wikipedia Articles.* [MSc Managmenet and Information Technology]. University of St Andrews.

Ashritha, K. S., Prajwala, T. M., & Chandrasekaran, K. (2017). Activity Theory Based Approach for Requirements Analysis of Android Applications. In L. Uden, W. Lu, & I.-H. Ting (Eds.), *Knowledge Management in Organizations* (pp. 3–15). Springer International Publishing. https://doi.org/10.1007/978-3-319-62698-7_1

Augustsson, D. (2021). Expansive Design for Teachers: An activity theoretical approach to design-based research. *EDeR. Educational Design Research*, *5*(1), Article 1. https://doi.org/10.15460/eder.5.1.1553

Baecker, R. M., Nastos, D., Posner, I. R., & Mawby, K. L. (1993). The user-centered iterative design of collaborative writing software. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '93*, 399–405.

Balestra, M., Zalmanson, L., Cheshire, C., Arazy, O., & Nov, O. (2017). It Was Fun, but Did It Last?: The Dynamic Interplay Between Fun Motives and Contributors' Activity in Peer

Production. *Proc. ACM Hum.-Comput. Interact.*, *1*(CSCW), 21:1-21:13. https://doi.org/10.1145/3134656

Barab, S. A., Evans, M. A., & Baek, E.-O. (2004). Activity Theory As a Lens for Characterizing the Participatory Unit. In *Handbook of research on educational communications and technology, 2nd ed* (pp. 199–214). Lawrence Erlbaum Associates Publishers.

Bear, J. B., & Collier, B. (2016). Where are the Women in Wikipedia? Understanding the Different Psychological Experiences of Men and Women in Wikipedia. *Sex Roles*, *74*(5), 254–265. https://doi.org/10.1007/s11199-015-0573-y

Beattie, A., Shaw, A., Kaur, S., & Kessler, D. (2009). Primary-care patients' expectations and experiences of online cognitive behavioural therapy for depression: A qualitative study. *Health Expectations: An International Journal of Public Participation in Health Care and Health Policy*, *12*(1), 45–59. https://doi.org/10.1111/j.1369-7625.2008.00531.x

Beck, E. (2002). P for Political: Participation is Not Enough. *Scandinavian Journal of Information Systems, 14*(1). https://aisel.aisnet.org/sjis/vol14/iss1/1

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? 🦜. *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610–623. https://doi.org/10.1145/3442188.3445922

Bergvall-Kåreborn, B., & Ståhlbrost, A. (2008). Participatory design: One step back or two steps forward? *Proceedings of the Tenth Anniversary Conference on Participatory Design 2008*, 102–111.

Bertelsen, O. W., & Bødker, S. (2003). Activity Theory. In *HCI models, theories, and frameworks: Toward a multidisciplinary science*. Morgan Kaufmann, c2003.

Bipat, T., McDonald, D. W., & Zachry, M. (2018). Do We All Talk Before We Type? Understanding Collaboration in Wikipedia Language Editions. *Proceedings of the 14th International Symposium on Open Collaboration*, 1–11. https://doi.org/10.1145/3233391.3233542

Biuk-Aghai, R. P., & Lei, K. H. (2010). Chatting in the Wiki: Synchronous-asynchronous integration. *Proceedings of the 6th International Symposium on Wikis and Open Collaboration*, 1–2. https://doi.org/10.1145/1832772.1832801

Bjork-James, C. (2021). New maps for an inclusive Wikipedia: Decolonial scholarship and strategies to counter systemic bias. *New Review of Hypermedia and Multimedia*, *27*(3), 207–228. https://doi.org/10.1080/13614568.2020.1865463

Black, L. W., Welser, H. T., Cosley, D., & DeGroot, J. M. (2011). Self-Governance Through Group Discussion in Wikipedia. *Small Group Research*, *42*(5), 595–634. https://doi.org/10.1177/1046496411406137

Blin, F., & Appel, C. (2011). Computer Supported Collaborative Writing in Practice: An Activity Theoretical Study. *CALICO Journal*, *28*(2), 473–497.

Bloomfield, D., & Nguyen, H. (2015). Creating and Sustaining Professional Learning Partnerships: Activity Theory as an Analytic Tool. *Australian Journal of Teacher Education*, *40*(11). https://doi.org/10.14221/ajte.2015v40n11.2

Bødker, S. (1990). Activity Theory as a Challenge to Systems Design. *DAIMI Report Series*, *334*, Article 334. https://doi.org/10.7146/dpb.v19i334.6564

Bødker, S., & Klokmose, C. N. (2011). The Human–Artifact Model: An Activity Theoretical Approach to Artifact Ecologies. *Human–Computer Interaction*, *26*(4), 315–371. https://doi.org/10.1080/07370024.2011.626709

Bødker, S., & Klokmose, C. N. (2012). Preparing Students for (Inter-)Action with Activity Theory. *International Journal of Design; Vol 6, No 3 (2012)*. http://www.ijdesign.org/index.php/IJDesign/article/view/972

Bødker, S., Lyle, P., & Saad-Sulonen, J. (2017). Untangling the Mess of Technological Artifacts: Investigating Community Artifact Ecologies. *Proceedings of the 8th International Conference on Communities and Technologies*, 246–255. https://doi.org/10.1145/3083671.3083675

Bonneau, C. (2013, July 6). *Contradictions and their concrete manifestations: An activity-theoretical analysis of the intra-organizational co-configuration of open source software*. Sub-theme 50: Activity Theory and Organizations — Proceedings of the 29th EGOS Colloquium (2013).

Botero, A., Hyysalo, S., Kohtala, C., & Whalen, J. (2020). Getting Participatory Design Done: From Methods and Choices to Translation Work across Constituent Domains. *International Journal of Design; Vol 14, No 2 (2020)*. http://www.ijdesign.org/index.php/IJDesign/article/view/3781

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Bryant, S. L., Forte, A., & Bruckman, A. (2005). Becoming Wikipedian: Transformation of participation in a collaborative online encyclopedia. *Proceedings of the 2005 International ACM SIGGROUP Conference on Supporting Group Work*, 1–10. https://doi.org/10.1145/1099203.1099205

Bryman, A. (2001). Qualitative Data Analysis. In *Social Research Methods*. Oxford University Press.

Callahan, E. S., & Herring, S. C. (2011). Cultural bias in Wikipedia content on famous persons. *Journal of the American Society for Information Science and Technology*, *62*(10), 1899–1915. https://doi.org/10.1002/asi.21577

Carvalho, P. de, & Fabiano, A. (2022). *Practising Thematic Analysis: From In-Depth Qualitative Data to Implications for Design*. https://doi.org/10.48340/ecscw2022_mc02

Category:Wikipedia feature requests—Wikipedia. (2022, August 14). Wikipedia. https://web.archive.org/web/20220814223200/https://en.wikipedia.org/wiki/Category:Wikipedia_feature_requests

Chełkowski, T., Gloor, P., & Jemielniak, D. (2016). Inequalities in Open Source Software Development: Analysis of Contributor's Commits in Apache Software Foundation Projects. *PLoS One*, *11*(4), e0152976. https://doi.org/10.1371/journal.pone.0152976

Chisik, Y., & Mancini, C. (2019). P for Politics D for Dialogue: Reflections on Participatory Design with Children and Animals. *Proceedings of the Sixth International Conference on Animal-Computer Interaction*, 1–13. https://doi.org/10.1145/3371049.3371061

Ciampaglia, G. L., & Taraborelli, D. (2015). MoodBar: Increasing New User Retention in Wikipedia through Lightweight Socialization. *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, 734–742. https://doi.org/10.1145/2675133.2675181

Ciffolilli, A. (2003). Phantom authority, self-selective recruitment and retention of members in virtual communities: The case of Wikipedia. *First Monday*. https://doi.org/10.5210/fm.v8i12.1108

Clark, T. (2008). `We're Over-Researched Here!': Exploring Accounts of Research Fatigue within Qualitative Research Engagements. *Sociology*, *42*(5), 953–970. https://doi.org/10.1177/0038038508094573

Clarke, K., Hughes, J., Martin, D., Rouncefield, M., Sommerville, I., Gurr, C., Hartswood, M., Procter, R., Slack, R., & Voss, A. (2003). Dependable Red Hot Action. In K. Kuutti, E. H. Karsten, G. Fitzpatrick, P. Dourish, & K. Schmidt (Eds.), *ECSCW 2003* (pp. 61–80). Springer Netherlands. https://doi.org/10.1007/978-94-010-0068-0_4

Clement, A., Costantino, T., Kurtz, D., & Tissenbaum, M. (2008). Participatory design and web 2.0: The case of PIPWatch, the collaborative privacy toolbar. *Proceedings of the Tenth Anniversary Conference on Participatory Design 2008*, 51–60. https://ojs.ruc.dk/index.php/pdc/article/view/1817

Collier, B., & Bear, J. (2012). Conflict, Criticism, or Confidence: An Empirical Examination of the Gender Gap in Wikipedia Contributions. *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work*, 383–392. https://doi.org/10.1145/2145204.2145265

Collins, P., Shukla, S., & Redmiles, D. (2002). Activity Theory and System Design: A View from the Trenches. *Computer Supported Cooperative Work (CSCW)*, *11*(1), 55–80. https://doi.org/10.1023/A:1015219918601

Community Insights/2018 Report/Contributors—Meta. (2023, January 9). Wikipedia. https://web.archive.org/web/20230109055837/https://meta.wikimedia.org/wiki/Community_Insights/2018_Report/Contributors#1._What_are_the_changes_of_diversity_of_participants_on_the_Wikimedia_projects?

Community Wishlist Survey 2016/Categories/Editing—Meta. (2016). https://web.archive.org/web/20220405110104/https://meta.wikimedia.org/wiki/Community_Wishlist_Survey_2016/Categories/Editing#Real_time_group_editing

Community Wishlist Survey 2020/Archive—Meta. (2020). https://web.archive.org/web/20230217140127/https://meta.wikimedia.org/wiki/Community_Wishlist_Survey_2020/Archive

Community Wishlist Survey/Help us—Meta. (2022). Wikipedia. https://web.archive.org/web/20220205020116/https://meta.wikimedia.org/wiki/Community_Wishlist_Survey/Help_us

Community Wishlist Survey—Meta. (2021). https://web.archive.org/web/20230317162358/https://meta.wikimedia.org/w/index.php?title=Community_Wishlist_Survey&oldid=22087455

Constantin, A., Alexandru, C., Korte, J., Wilson, C., Fails, J. A., Sim, G., Read, J. C., & Eriksson, E. (2021). Distributing participation in design: Addressing challenges of a global pandemic. *International Journal of Child-Computer Interaction*, *28*, 100255. https://doi.org/10.1016/j.ijcci.2021.100255

Content translation/Machine Translation/Google Translate—MediaWiki. (2022, November 30). https://web.archive.org/web/20230130023859/https://www.mediawiki.org/wiki/Content_translation/Machine_Translation/Google_Translate#Can_this_content_be_used_for_improving_machine_translation_systems_in_general

Content translation—MediaWiki. (2023, January 21). MediaWiki. https://web.archive.org/web/20230121212735/https://www.mediawiki.org/wiki/Content_translation

Crabtree, A. (1998). Ethnography in Participatory Design. *Proceedings of the 1998 Participatory Design Conference*. https://eprints.lancs.ac.uk/id/eprint/11565/

Crabtree, A., Rouncefield, M., & Tolmie, P. (2012a). Design Ethnography in a Nutshell. In *Doing Design Ethnography* (pp. 183–205). Springer. https://doi.org/10.1007/978-1-4471-2726-0_10

Crabtree, A., Rouncefield, M., & Tolmie, P. (2012b). Ethnography and Systems Design. In *Doing Design Ethnography* (pp. 7–21). Springer. https://doi.org/10.1007/978-1-4471-2726-0_10

Crowston, K., & Rezgui, A. (2020). Effects of Stigmergic and Explicit Coordination on Wikipedia Article Quality. *HICSS*, 1–10. https://doi.org/10.24251/HICSS.2020.287

Cummings, R. E. (2020). Writing knowledge: Wikipedia, public review, and peer review. *Studies in Higher Education*, *45*(5), 950–962. https://doi.org/10.1080/03075079.2020.1749791

Dahl, Y., & Sharma, K. (2022). Six Facets of Facilitation: Participatory Design Facilitators' Perspectives on Their Role and Its Realization. *CHI Conference on Human Factors in Computing Systems*, 1–14. https://doi.org/10.1145/3491102.3502013

D'Angelo, G., Di Iorio, A., & Zacchiroli, S. (2018). Spacetime Characterization of Real-Time Collaborative Editing. *Proc. ACM Hum.-Comput. Interact.*, *2*(CSCW), 41:1-41:19. https://doi.org/10.1145/3274310

Danielsson, K., Naghsh, A. M., Gumm, D., & Warr, A. (2008). Distributed participatory design. *CHI '08 Extended Abstracts on Human Factors in Computing Systems*, 3953–3956. https://doi.org/10.1145/1358628.1358965

Decker, B., Ras, E., Rech, J., Jaubert, P., & Rieth, M. (2007). Wiki-Based Stakeholder Participation in Requirements Engineering. *IEEE Software*, *24*(2), 28–35. https://doi.org/10.1109/MS.2007.60

Design/Statement of purpose—MediaWiki. (2016). https://web.archive.org/web/20221103124650/https://www.mediawiki.org/wiki/Design/Statement_of_purpose

Dimitrakopoulos, G., & Uden, L. (2020). Chapter 4—User requirements and preferences for ITS. In G. Dimitrakopoulos, L. Uden, & I. Varlamis (Eds.), *The Future of Intelligent Transport Systems* (pp. 43–62). Elsevier. https://doi.org/10.1016/B978-0-12-818281-9.00004-8

Dourish, P. (2006). Implications for design. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 541–550. https://doi.org/10.1145/1124772.1124855

Ducheneaut, N. (2005). Socialization in an Open Source Software Community: A Socio-Technical Analysis. *Computer Supported Cooperative Work (CSCW)*, *14*(4), 323–368. https://doi.org/10.1007/s10606-005-9000-1

Dudley, J. J., Jacques, J. T., & Kristensson, P. O. (2019). Crowdsourcing Interface Feature Design with Bayesian Optimization. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–12. https://doi.org/10.1145/3290605.3300482

Durst, C., Wiser, F., Luber, S., & Wickramasinghe, N. (2020). Development of an activity theory-based framework for the analysis and design of socio-technical systems. *International Journal of Networking and Virtual Organisations*, *23*(3), 261. https://doi.org/10.1504/IJNVO.2020.109322

Ebersbach, A., Glaser, M., Heigl, R., & Warta, A. (Eds.). (2008). The Wiki Concept. In *Wiki: Web Collaboration* (pp. 11–34). Springer. https://doi.org/10.1007/978-3-540-68173-1_2

Eckert, S., & Steiner, L. (2013). Wikipedia's gender gap. *Media Disparity: A Gender Battleground*, 87–98.

Edwards, P. N., Jackson, S. J., Bowker, G. C., & Knobel, C. P. (2007). *Understanding Infrastructure: Dynamics, Tensions, and Design* [Working Paper]. http://deepblue.lib.umich.edu/handle/2027.42/49353

Ellis, C. A., Gibbs, S. J., & Rein, G. (1991). Groupware: Some Issues and Experiences. *Communications of the ACM, 34*(1), 39–58. https://doi.org/10.1145/99977.99987

Engeström, Y. (1987). *Learning by expanding: An activity-theoretical approach to developmental research* [Doctoral Thesis]. Orienta-konsultit.

Engeström, Y. (2001). Expansive Learning at Work: Toward an activity theoretical reconceptualization. *Journal of Education and Work, 14*(1), 133–156. https://doi.org/10.1080/13639080020028747

Engeström, Y. (Ed.). (2014). The Emergence of Learning Activity as a Historical Form of Human Learning. In *Learning by Expanding: An Activity-Theoretical Approach to Developmental Research* (2nd ed., pp. 25–108). Cambridge University Press. https://doi.org/10.1017/CBO9781139814744.004

English Wikipedia's Stub Articles. (2021). https://web.archive.org/web/20210531135336/https://en.wikipedia.org/wiki/Category:All_stub_articles

Erhart, E., & Halfaker, A. (2015, September 25). Wikipedia's very active editor numbers have stabilized—Delve into the data with us. *Diff*. https://diff.wikimedia.org/2015/09/25/wikipedia-editor-numbers/

Exter, K., Rowe, S., Boyd, W., & Lloyd, D. (2012). Using Web 2.0 Technologies for Collaborative Learning in Distance Education—Case Studies from an Australian University. *Future Internet, 4*, 216–237. https://doi.org/10.3390/fi4010216

Farzan, R., Savage, S., & Saviaga, C. F. (2016). Bring on Board New Enthusiasts! A Case Study of Impact of Wikipedia Art + Feminism Edit-A-Thon Events on Newcomers. In E. Spiro & Y.-Y. Ahn (Eds.), *Social Informatics* (pp. 24–40). Springer International Publishing. https://doi.org/10.1007/978-3-319-47880-7_2

Foot, K. A. (2014). Cultural-Historical Activity Theory: Exploring a Theory to Inform Practice and Research. *Journal of Human Behavior in the Social Environment*, *24*(3), 329–347. https://doi.org/10.1080/10911359.2013.831011

Ford, H. (2017). The Search for Wikipedia's Edges. In *The Routledge Companion to Digital Ethnography* (pp. 442–451). Routledge.

Ford, H., & Graham, M. (2015). *Semantic Cities: Coded Geopolitics and the Rise of the Semantic Web* (SSRN Scholarly Paper 2682459). https://papers.ssrn.com/abstract=2682459

Ford, H., Graham, M., & Meyer, E. (2015). *Fact factories: Wikipedia and the power to represent* [Http://purl.org/dc/dcmitype/Text, University of Oxford]. https://ora.ox.ac.uk/objects/uuid:b34fdd6c-ec15-4bcd-acba-66a777739b4d

Ford, H., & Wajcman, J. (2017). 'Anyone can edit', not everyone does: Wikipedia's infrastructure and the gender gap. *Social Studies of Science*, *47*(4), 511–527. https://doi.org/10.1177/0306312717692172

Gardner, S. (2011, February 20). Nine Reasons Women Don't Edit Wikipedia (in their own words). *Sue Gardner's Blog*. https://suegardner.org/2011/02/19/nine-reasons-why-women-dont-edit-wikipedia-in-their-own-words/

Gazan, R. (2011). Redesign as an act of violence: Disrupted interaction patterns and the fragmenting of a social Q&A community. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2847–2856. https://doi.org/10.1145/1978942.1979365

Geiger, R. S., & Halfaker, A. (2013). Using Edit Sessions to Measure Participation in Wikipedia. *Proceedings of the 2013 Conference on Computer Supported Cooperative Work*, 861–870. https://doi.org/10.1145/2441776.2441873

Gender bias on Wikipedia—Wikipedia. (2023, January 28). https://web.archive.org/web/20230128042356/https://en.wikipedia.org/wiki/Gender_bias_on_Wikipedia

Georg, G. (2011). *Activity Theory and its Applications in Software Engineering and Technology* [Technical Report CS-11-101 Literature Search Results and Observations]. Colorado State University. https://www.cs.colostate.edu/TechReports/Reports/2011/tr11-101.pdf

Gheitasy, A. (2017). *Socio-technical gaps and social capital formation in Online Collaborative Consumption communities* [PhD Thesis, University of West London]. http://repository.uwl.ac.uk/id/eprint/3835/

Given, L. (2008). *The SAGE Encyclopedia of Qualitative Research Methods*. SAGE Publications, Inc. https://doi.org/10.4135/9781412963909

Gluza, W., Turaj, I., & Meier, F. (2021). Wikipedia Edit-a-thons and Editor Experience: Lessons from a Participatory Observation. *17th International Symposium on Open Collaboration*, 1–9. https://doi.org/10.1145/3479986.3479994

Goldman, E. (2009). *Wikipedia's Labor Squeeze and its Consequences*.

Gonçalves, A., Sousa, P., & Zacarias, M. (2013). A Method to Capture Activity Diagrams from Ontological Model. *Procedia Technology*, *9*, 553–562. https://doi.org/10.1016/j.protcy.2013.12.062

Good, A., & Omisade, O. (2019). Linking Activity Theory with User Centred Design: A Human Computer Interaction Framework for the Design and Evaluation of mHealth Interventions. *Studies in Health Technology and Informatics*, *263*, 49–63. https://doi.org/10.3233/SHTI190110

Graham, M. (2011). Wiki Space: Palimpsests and the Politics of Exclusion. In G. Lovink & N. Tkacz (Eds.), *Critical Point of View: A Wikipedia Reader: Vol. Amsterdam: Institute of Network Cultures* (pp. 269–282). Institute of Network Cultures.

Graham, M. (2014). *The Knowledge Based Economy and Digital Divisions of Labour* (SSRN Scholarly Paper 2363880). https://papers.ssrn.com/abstract=2363880

Graham, M., & Hogan, B. (2014). *Uneven Openness: Barriers to MENA Representation on Wikipedia* (SSRN Scholarly Paper ID 2430912). Social Science Research Network. https://papers.ssrn.com/abstract=2430912

Graham, M., Hogan, B., Straumann, R. K., & Medhat, A. (2014). Uneven Geographies of User-Generated Information: Patterns of Increasing Informational Poverty. *Annals of the Association of American Geographers*, *104*(4), 746–764. https://doi.org/10.1080/00045608.2014.910087

Graham, M., Straumann, R. K., & Hogan, B. (2015). Digital Divisions of Labor and Informational Magnetism: Mapping Participation in Wikipedia. *Annals of the Association*

of American Geographers, *105*(6), 1158–1178. https://doi.org/10.1080/00045608.2015.1072791

Grant, K. D. (2001). Dimensions of collaborative work. *Human Computer Interaction Consortium 2001, Colorado*, 7–11.

Greenbaum, J., & Halskov, K. (1993). PD a personal statement. *Communications of the ACM*, *36*(6), 47. https://doi.org/10.1145/153571.214816

Greenbaum, J., & Loi, D. (2012). Participation, the camel and the elephant of design: An introduction. *CoDesign*, *8*(2–3), 81–85. https://doi.org/10.1080/15710882.2012.690232

Grudin, J. (1990). The computer reaches out: The historical continuity of interface design. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 261–268. https://doi.org/10.1145/97243.97284

Grudin, J., & Pruitt, J. (2002). Personas, Participatory Design and Product Development: An Infrastructure for Engagement. *Proceedings of Participation and Design Conference (PDC2002), Sweden*, 144–161. http://www.itee.uq.edu.au/~comp4501/_2003/_Readings/GrudinPersonas.pdf

Hagen, P., Collin, P., Metcalf, A., Nicholas, M., Rahilly, K., & Swainston, N. (2012). *Participatory design of evidence-based online youth mental health promotion, intervention and treatment* [Research report]. Young and Well Cooperative Research Centre - Western Sydney University. https://researchdirect.westernsydney.edu.au/islandora/object/uws%3A18814/

Halfaker, A., Geiger, R. S., Morgan, J. T., & Riedl, J. (2013). The Rise and Decline of an Open Collaboration System: How Wikipedia's Reaction to Popularity Is Causing Its Decline. *American Behavioral Scientist*, *57*(5), 664–688. https://doi.org/10.1177/0002764212469365

Halfaker, A., Kittur, A., & Riedl, J. (2011). Don't bite the newbies: How reverts affect the quantity and quality of Wikipedia work. *Proceedings of the 7th International Symposium on Wikis and Open Collaboration*, 163–172. https://doi.org/10.1145/2038558.2038585

Hammarberg, K., Kirkman, M., & de Lacey, S. (2016). Qualitative research methods: When to use them and how to judge them. *Human Reproduction*, *31*(3), 498–501. https://doi.org/10.1093/humrep/dev334

Hartson, R., & Pyla, P. S. (2012). Chapter 4 - Contextual Analysis: Consolidating and Interpreting Work Activity Data. In R. Hartson & P. S. Pyla (Eds.), *The UX Book* (pp. 129–160). Morgan Kaufmann. https://doi.org/10.1016/B978-0-12-385241-0.00004-X

Hartswood, M., Procter, R., Slack, R., Voß, A., Büscher, M., Rouncefield, M., & Rouchy, P. (2008). Co-Realization: Toward a Principled Synthesis of Ethnomethodology and Participatory Design. In M. S. Ackerman, C. A. Halverson, T. Erickson, & W. A. Kellogg (Eds.), *Resources, Co-Evolution and Artifacts: Theory in CSCW* (pp. 59–94). Springer. https://doi.org/10.1007/978-1-84628-901-9_3

Hasan, H. (2003, January 6). An activity-based model of collective knowledge. *Faculty of Commerce - Papers (Archive)*. Annual Hawaii International Conference on System Sciences (HICSS), Big Island, HI, USA. https://doi.org/10.1109/HICSS.2003.1174230

Hautasaari, A. (2013). 'Could Someone Please Translate This?': Activity Analysis of Wikipedia Article Translation by Non-Experts. *Proceedings of the 2013 Conference on Computer Supported Cooperative Work*, 945–954. https://doi.org/10.1145/2441776.2441883

Heintz, M., Law, E., Govaerts, S., Holzer, A., & Gillet, D. (2014). *Pdot: Participatory design online tool*. https://doi.org/10.1145/2559206.2581266

Heintz, M. M. (2017). *Software-supported participatory design: Design and evaluation of the tool PDot* [Ph.D., University of Leicester]. https://figshare.com/articles/Software-supported_Participatory_Design_Design_and_Evaluation_of_the_Tool_PDot/10205576

Heiskanen, E., Hyysalo, S., Kotro, T., & Repo, P. (2010). Constructing innovative users and user-inclusive innovation communities. *Technology Analysis & Strategic Management*, *22*(4), 495–511. https://doi.org/10.1080/09537321003714568

Help:Content translation/Translating/Initial machine translation—MediaWiki. (2022). https://web.archive.org/web/20221230010747/https://www.mediawiki.org/wiki/Help:Content_translation/Translating/Initial_machine_translation#Machine_translation_availability

Help:Edit conflict. (2022). In *Wikipedia*. https://web.archive.org/web/20220430225120/https://en.wikipedia.org/wiki/Help:Edit_conflict

Hern, A. (2016, February 26). Head of Wikimedia resigns over search engine plans. *The Guardian*.

https://web.archive.org/web/20220422214649/https://www.theguardian.com/technology/2016/feb/26/wikimedia-head-lila-tretikov-resigns-search-engine-plans

Hewitt, J. (2004). An Exploration of Community in a Knowledge Forum Classroom: An Activity System Analysis. In J. H. Gray, R. Kling, & S. Barab (Eds.), *Designing for Virtual Communities in the Service of Learning* (pp. 210–238). Cambridge University Press. https://doi.org/10.1017/CBO9780511805080.012

Heylighen, F. (2006). Why is Open Access Development so Successful? Stigmergic organization and the economics of information. *ArXiv:Cs/0612071*, *abs/cs/0612071*. https://doi.org/10.48550/arXiv.cs/0612071

Hine, C. (2000). *Virtual Ethnography*. SAGE Publications Ltd. https://doi.org/10.4135/9780857020277

Hine, C. (2015). *Ethnography for the Internet: Embedded, Embodied and Everyday*. Routledge. https://doi.org/10.4324/9781003085348

Ho, R. (2021, June 15). Nurturing newbies: Growing Wikipedia contributors by design. *Medium - Wikimedia Design*. https://web.archive.org/web/20210615111551/https://uxdesign.cc/jun-12-2021-9aa6bedcee02?gi=8a5e49aec1af

Hube, C. (2017). Bias in Wikipedia. *Proceedings of the 26th International Conference on World Wide Web Companion*, 717–721. https://doi.org/10.1145/3041021.3053375

Hudson, J. (2018). Using Wikis for Collaborative Writing in the ELT Classroom. *International Journal of Pedagogy and Teacher Education*, *2*, 413–426. https://doi.org/10.20961/ijpte.v2i2.22906

Hughes, J. A., Randall, D., & Shapiro, D. (1992a). Faltering from ethnography to design. *Proceedings of the 1992 ACM Conference on Computer-Supported Cooperative Work*, 115–122. https://doi.org/10.1145/143457.143469

Hughes, J. A., Randall, D., & Shapiro, D. (1992b). From ethnographic record to system design: Some experiences from the field. *Computer Supported Cooperative Work (CSCW)*, *1*(3), 123–141. https://doi.org/10.1007/BF00752435

Huybrechts, L., Hendriks, N., Yndigegn, S. L., & Malmborg, L. (2018). Scripting: An exploration of designing for participation over time with communities. *CoDesign*, *14*(1), 17–31. https://doi.org/10.1080/15710882.2018.1424205

Hyysalo, S. (2000). *Representing and constructing use—An activity theoretical perpective to design work | PDC*. https://ojs.ruc.dk/index.php/pdc/article/view/221

Hyysalo, S. (2005). Objects and Motives in a Product Design Process. *Mind, Culture, and Activity*, *12*(1), 19–36. https://doi.org/10.1207/s15327884mca1201_3

Hyysalo, S., Jensen, T. E., & Oudshoorn, N. (2016). Introduction to the New Production of Users. In *The New Production of Users*. Routledge.

Ilyenkov, E. (1982). The Concrete and the Dialectics of the Universal and the Individual. In *Dialectics of the Abstract & the Concrete in Marx's Capital*. https://www.marxists.org/archive/ilyenkov/works/abstract/index.htm

Jemielniak, D. (2012). *Trust, Control, and Formalization in Open-Collaboration Communities: A Qualitative Study of Wikipedia*. https://depot.ceon.pl/handle/123456789/308

Jemielniak, D. (2016). Breaking the glass ceiling on Wikipedia. *Feminist Review*, *113*(1), 103–108. https://doi.org/10.1057/fr.2016.9

Jemielniak, D., & Przegalinska, A. (2020). Peer Production. In *Collaborative Society* (pp. 41–59). MIT Press. https://ieeexplore.ieee.org/document/9085804

Jemielniak, D., & Raburski, T. (2014). Liquid collaboration? In *Liquid Organization* (pp. 85–103). Routledge. https://crow.kozminski.edu.pl/papers/Liquid3.pdf

Jemielniak, D., & Wilamowski, M. (2017). Cultural diversity of quality of information on Wikipedias. *Journal of the Association for Information Science and Technology*.

Jimmy Wales (Director). (2007, January 16). *Jimmy Wales: How a ragtag band created Wikipedia - YouTube - TED Talk*. https://youtu.be/WQR0gx0QBZ4?t=678

Johansen, R. (1988). *GroupWare: Computer Support for Business Teams*. The Free Press.

Johnson, L. L. (2016). Writing 2.0: How English Teachers Conceptualize Writing with Digital Technologies. *English Education*, *49*(1), 28–62.

Jonassen, D. H., & Rohrer-Murphy, L. (1999). Activity theory as a framework for designing constructivist learning environments. *Educational Technology Research and Development*, *47*(1), 61–79. https://doi.org/10.1007/BF02299477

Judge, T., Matthews, T., & Whittaker, S. (2012). Comparing collaboration and individual personas for the design and evaluation of collaboration software. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1997–2000. https://doi.org/10.1145/2207676.2208344

Kaptelinin, V., Nardi, B. A., & Macaulay, C. (1999). Methods & tools: The activity checklist: a tool for representing the "space" of context. *Interactions*, *6*(4), 27–39. https://doi.org/10.1145/306412.306431

Karanasios, S., Riisla, K., & Simeonova, B. (2017). *Exploring the use of contradictions in activity theory studies: An interdisciplinary review*. https://repository.lboro.ac.uk/articles/conference_contribution/Exploring_the_use_of_contradictions_in_activity_theory_studies_An_interdisciplinary_review/9499139/1

Kautz, K., Bjerknes, G., Fisher, J., & Jensen, T. (2018). *Distributed Participatory Design in Crowdsourced Information Systems Development*. https://doi.org/10.5130/acis2018.aa

Kazman, R., & Chen, H.-M. (2009). The metropolis model a new logic for development of crowdsourced systems. *Communications of the ACM*, *52*(7), 76–84. https://doi.org/10.1145/1538788.1538808

Kensing, F., & Greenbaum, J. (2013). Heritage: Having a say. In *Routledge International Handbook of Participatory Design*. Routledge.

Kimmons, R. M. (2011). Understanding collaboration in Wikipedia. *First Monday*, *16*(12). https://doi.org/10.5210/fm.v16i12.3613

Kittur, A., Chi, E., Pendleton, B. A., & Suh, B. (2006). *Power of the Few vs. Wisdom of the Crowd: Wikipedia and the Rise of the Bourgeoisie*.

Kittur, A., & Kraut, R. E. (2008). Harnessing the Wisdom of Crowds in Wikipedia: Quality through Coordination. *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work*, 37–46. https://doi.org/10.1145/1460563.1460572

Kline, R. R. (2002). *Consumers in the Country: Technology and Social Change in Rural America*. https://doi.org/10.56021/9780801862489

Koch, B., Denton, E., Hanna, A., & Foster, J. G. (2021). *Reduced, Reused and Recycled: The Life of a Dataset in Machine Learning Research* (arXiv:2112.01716). arXiv. https://doi.org/10.48550/arXiv.2112.01716

Konieczny, P. (2016). Teaching with Wikipedia in a 21st-Century Classroom: Perceptions of Wikipedia and Its Educational Benefits. *Journal of the Association for Information Science and Technology*, *67*(7), 1523–1534. https://doi.org/10.1002/asi.23616

Korsgaard, H., Klokmose, C. N., & Bødker, S. (2016). Computational alternatives in participatory design: Putting the t back in socio-technical research. *Proceedings of the*

14th Participatory Design Conference: Full Papers - Volume 1, 71–79. https://doi.org/10.1145/2940299.2940314

Kothari, C. (2004). Research Methodology: An Introduction. In *Research Mathodology: Methods and Techniques* (pp. 1–21). Published by New Age International (P) Ltd., Publishers. http://www.modares.ac.ir/uploads/Agr.Oth.Lib.17.pdf

Kriplean, T., Beschastnikh, I., & McDonald, D. W. (2008). Articulations of Wikiwork: Uncovering Valued Work in Wikipedia through Barnstars. *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work*, 47–56. https://doi.org/10.1145/1460563.1460573

Kriplean, T., Beschastnikh, I., McDonald, D. W., & Golder, S. A. (2007). Community, consensus, coercion, control: Cs*w or how policy mediates mass participation. *Proceedings of the 2007 ACM International Conference on Supporting Group Work*, 167–176. https://doi.org/10.1145/1316624.1316648

Krueger, R. A., & Casey, M. A. (2014). *Focus Groups: A Practical Guide for Applied Research* (Fifth edition). SAGE Publications, Inc.

Kumar, S., West, R., & Leskovec, J. (2016). Disinformation on the Web: Impact, Characteristics, and Detection of Wikipedia Hoaxes. *Proceedings of the 25th International Conference on World Wide Web*, 591–602.

Kupferschmidt, K. (2022). As Musk reshapes Twitter, academics ponder taking flight. *Science.Org*, *378*(6620). https://www.science.org/content/article/musk-reshapes-twitter-academics-ponder-taking-flight

Kuutti, K. (1991). The concept of activity as a basic unit of analysis for CSCW research. In L. Bannon, M. Robinson, & K. Schmidt (Eds.), *Proceedings of the Second European Conference on Computer-Supported Cooperative Work ECSCW '91* (pp. 249–264). Springer Netherlands. https://doi.org/10.1007/978-94-011-3506-1_19

Kuutti, K. (1995). Activity theory as a potential framework for human-computer interaction research. In *Context and consciousness: Activity theory and human-computer interaction* (pp. 17–44). Massachusetts Institute of Technology.

Kuznetsov, S. (2006). Motivations of contributors to Wikipedia. *ACM SIGCAS Computers and Society*, *36*(2), 1-es. https://doi.org/10.1145/1215942.1215943

Lam, S. (Tony) K., Uduwage, A., Dong, Z., Sen, S., Musicant, D. R., Terveen, L., & Riedl, J. (2011). WP:clubhouse? An exploration of Wikipedia's gender imbalance. *Proceedings of the 7th International Symposium on Wikis and Open Collaboration*, 1–10. https://doi.org/10.1145/2038558.2038560

Lampe, C., Obar, J., Ozkaya, E., Zube, P., & Velasquez, A. (2012). Classroom Wikipedia Participation Effects on Future Intentions to Contribute. *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work*, 403–406. https://doi.org/10.1145/2145204.2145267

Langrock, I., & González-Bailón, S. (2022). The Gender Divide in Wikipedia: Quantifying and Assessing the Impact of Two Feminist Interventions. *Journal of Communication*, *72*(3), 297–321. https://doi.org/10.1093/joc/jqac004

Larsen-Ledet, I., Korsgaard, H., & Bødker, S. (2020). Collaborative Writing Across Multiple Artifact Ecologies. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–14. https://doi.org/10.1145/3313831.3376422

Larusson, J. A., & Alterman, R. (2009). Wikis to support the "collaborative" part of collaborative learning. *International Journal of Computer-Supported Collaborative Learning*, *4*(4), 371–402. https://doi.org/10.1007/s11412-009-9076-6

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation* (p. 138). Cambridge University Press. https://doi.org/10.1017/CBO9780511815355

Lazar, J., Feng, J. H., & Hochheiser, H. (2017). *Research methods in human-computer interaction*.

Lazar, Jonathan., Feng, J. Heidi., & Hochheiser, Harry. (2010). Ethnography. In *Research methods in human-computer interaction* (pp. 142–155). Wiley. http://library.st-andrews.ac.uk/search S5?/Yresearch+methods+Lazar&searchscope=5&SORT=DZ/Yresearch+methods+Lazar&searchscope=5&SORT=DZ&extended=0&SUBKEY=research+methods+Lazar/1%2C5%2C5%2CB/frameset&FF=Yresearch+methods+Lazar&searchscope=5&SORT=DZ&2%2C2%2C

Leigh, D. (2009). SWOT Analysis. In *Handbook of Improving Performance in the Workplace: Volumes 1-3* (pp. 115–140). John Wiley & Sons, Ltd. https://doi.org/10.1002/9780470592663.ch24

Leont′ev, A. N. (1978). *Activity, Consciousness, and Personality*.

Leskovec, J., Huttenlocher, D. P., & Kleinberg, J. M. (2010). Governance in Social Media: A case study of the Wikipedia promotion process. *CoRR, abs/1004.3547*. http://arxiv.org/abs/1004.3547

Li, A., & Farzan, R. (2018). Keeping up on Current Events! A Case Study of Newcomers to Wikipedia. In S. Staab, O. Koltsova, & D. I. Ignatov (Eds.), *Social Informatics* (pp. 348–369). Springer International Publishing.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Sage Publications.

List of Wikipedia controversies. (2022). In *Wikipedia*. https://web.archive.org/web/20221230074113/https://en.wikipedia.org/wiki/List_of_Wikipedia_controversies

List of Wikipedias—Meta. (2023). https://web.archive.org/web/20230107230014/https://meta.wikimedia.org/wiki/List_of_Wikipedias

Littlejohn, A., Hood, N., Rehm, M., McGill, L., Rienties, B., & Highton, M. (2019). Learning to become an online editor: The editathon as a learning environment. *Interactive Learning Environments*, *0*(0), 1–14. https://doi.org/10.1080/10494820.2019.1625557

Liu, F., Lin, Z., & Qu, Y. (2020). System Dynamics Modeling for Online Encyclopedias Based on Activity Theory. *Proceedings of the 2020 The 6th International Conference on Frontiers of Educational Technologies*, 173–177. https://doi.org/10.1145/3404709.3404753

Liu, J., & Ram, S. (2011). Who does what: Collaboration patterns in the wikipedia and their impact on article quality. *ACM Transactions on Management Information Systems*, *2*(2), 1–23. https://doi.org/10.1145/1985347.1985352

Löwgren, J. (1995). Applying design methodology to software development. *Proceedings of the 1st Conference on Designing Interactive Systems: Processes, Practices, Methods, & Techniques*, 87–95. https://doi.org/10.1145/225434.225444

Lukyanenko, R., Parsons, J., Wiersma, Y., Sieber, R., & Maddah, M. (2016). Participatory Design for User-generated Content: Understanding the challenges and moving forward. *Scandinavian Journal of Information Systems*, *28*(1). https://aisel.aisnet.org/sjis/vol28/iss1/2

Ma, M.-Y., Wu, F.-G., & Chang, R.-H. (2007). A new design approach of user-centered design on a personal assistive bathing device for hemiplegia. *Disability and Rehabilitation*, *29*(14), 1077–1089. https://doi.org/10.1080/09638280600949712

Markel, J. M., & Guo, P. J. (2020). *Designing the Future of Experiential Learning Environments for a Post-COVID World: A Preliminary Case Study*. https://www.microsoft.com/en-us/research/publication/designing-the-future-of-experiential-learning-environments-for-a-post-covid-world-a-preliminary-case-study/

Marshall, G. (Ed.). (1994). *The Concise Oxford Dictionary of Sociology*. Oxford Paperbacks.

Mathers, C. D., Vos, E. T., Stevenson, C. E., & Begg, S. J. (2000). The Australian Burden of Disease Study: Measuring the loss of health from diseases, injuries and risk factors. *The Medical Journal of Australia*, *172*(12), 592–596. https://doi.org/10.5694/j.1326-5377.2000.tb124125.x

Matthews, T., Whittaker, S., Moran, T., & Yang, M. (2010). Collaboration personas: A framework for understanding designing collaborative workplace tools. *Workshop on Collective Intelligence in Organizations, ACM CSCW*.

Matthews, T., Whittaker, S., Moran, T., & Yuen, S. (2011). Collaboration personas: A new approach to designing workplace collaboration tools. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2247–2256. https://doi.org/10.1145/1978942.1979272

mbmusablog. (2018, April 12). A history of women in St Andrews: The Lady Literate in Arts scheme. *Museum Collections Blog*. https://museumoftheuniversityofstandrews.wordpress.com/2018/04/12/a-history-of-women-in-st-andrews-the-lady-literate-in-arts-scheme/

McCalla, L. E. (2019). *Activity, Creativity, and Design: Modeling and Assessing Student Development in a Project-Based Course* [Ph.D., University of Georgia]. https://www.proquest.com/pqdtglobal/docview/2356062287/abstract/156478DD1C7E41E7PQ/8

Mediawiki, T, E, Officer, O. of the C. E., Advancement, Communications, Administration, F. and, Legal, Operations, Product, Culture, T. and, & Technology. (2022). *Community Tech—Meta*. https://meta.wikimedia.org/wiki/Community_Tech

Mendoza-Chapa, S., Romero-Salcedo, M., & Oktaba, H. (2000). Group awareness support in collaborative writing systems. *Proceedings Sixth International Workshop on Groupware. CRIWG 2000*, 112–118. https://doi.org/10.1109/CRIWG.2000.885162

Menking, A., & Rosenberg, J. (2021). WP:NOT, WP:NPOV, and Other Stories Wikipedia Tells Us: A Feminist Critique of Wikipedia's Epistemology. *Science, Technology & Human Values*, *46*(3), 455–479.

Meuer, M., Nestler, S., & Oeberst, A. (2022). What determines hindsight bias in written work? One field and three experimental studies in the context of Wikipedia. *Journal of Experimental Psychology: Applied*, No Pagination Specified-No Pagination Specified. https://doi.org/10.1037/xap0000445

Miettinen, R., & Hasu, M. (2002). Articulating User Needs in Collaborative Design: Towards an Activity-Theoretical Approach. *Computer Supported Cooperative Work (CSCW)*, *11*(1), 129–151. https://doi.org/10.1023/A:1015256909032

Morgan, J. T., Bouterse, S., Walls, H., & Stierch, S. (2013). Tea and Sympathy: Crafting Positive New User Experiences on Wikipedia. *Proceedings of the 2013 Conference on Computer Supported Cooperative Work*, 839–848. https://doi.org/10.1145/2441776.2441871

Morgan, J. T., & Halfaker, A. (2018). Evaluating the impact of the Wikipedia Teahouse on newcomer socialization and retention. *Proceedings of the 14th International Symposium on Open Collaboration*, 1–7. https://doi.org/10.1145/3233391.3233544

Morgan, J. T., Mason, R. M., & Nahon, K. (2012). Negotiating Cultural Values in Social Media: A Case Study from Wikipedia. *2012 45th Hawaii International Conference on System Sciences*, 3490–3499. https://doi.org/10.1109/HICSS.2012.443

Movement Strategy/About—Meta. (2020). https://web.archive.org/web/20220930154705/https://meta.wikimedia.org/wiki/Movement_Strategy/About

Muller, M., & Druin, A. (2002). Participatory Design: The Third Space in HCI. *Handbook of HCI*.

Müller-Birn, C., Dobusch, L., & Herbsleb, J. D. (2013). Work-to-rule: The emergence of algorithmic governance in Wikipedia. *Proceedings of the 6th International Conference on Communities and Technologies*, 80–89. https://doi.org/10.1145/2482991.2482999

Murray, H. (2019). More than 2 billion pairs of eyeballs: Why aren't you sharing medical knowledge on Wikipedia? *BMJ Evid Based Med*, *24*(3), 90–91.

Musicant, D. R., Ren, Y., Johnson, J. A., & Riedl, J. (2011). Mentoring in Wikipedia: A clash of cultures. *Proceedings of the 7th International Symposium on Wikis and Open Collaboration*, 173–182.

Mwanza, D. (2001). *Where Theory meets Practice: A Case for an Activity Theory based Methodology to guide Computer System Design*. Proceedings of INTERACT' 2001: Eighth IFIP TC 13 Conference on Human-Computer Interaction, Tokyo, Japan. http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.17.6630

Nakamura, A., Suzuki, Y., & Ishikawa, Y. (2013). Clustering Editors of Wikipedia by Editor's Biases. *Proceedings of the 2013 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT) - Volume 01*, *1*, 351–358. https://doi.org/10.1109/WI-IAT.2013.50

Narayan, S., Orlowitz, J., Morgan, J., Hill, B. M., & Shaw, A. (2017). The Wikipedia Adventure: Field Evaluation of an Interactive Tutorial for New Users. *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, 1785–1799.

Nardi, B. A. (Ed.). (1995). Activity Theory and Human-Computer Interaction. In *Context and Consciousness: Activity Theory and Human-Computer Interaction* (p. 0). Massachusetts Institute of Technology. https://doi.org/10.7551/mitpress/2137.003.0005

Niederer, S., & van Dijck, J. (2010). Wisdom of the crowd or technicity of content? Wikipedia as a sociotechnical system. *New Media and Society*, *12*(8), 1368–1387. https://doi.org/10.1177/1461444810365297

Nielsen, J. (1994). Enhancing the explanatory power of usability heuristics. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 152–158. https://doi.org/10.1145/191666.191729

Noël, S., & Robert, J.-M. (2003). How the Web is used to support collaborative writing. *Behaviour & Information Technology*, *22*(4), 245–262. https://doi.org/10.1080/0144929031000120860

Noël, S., & Robert, J.-M. (2004). Empirical Study on Collaborative Writing: What Do Co-authors Do, Use, and Like? *Computer Supported Cooperative Work (CSCW)*, *13*(1), 63–89. https://doi.org/10.1023/B:COSU.0000014876.96003.be

Obendorf, H., & Janneck, de M. (2009). *Inter-Contextual Distributed Participatory Design Communicating design philosophy and enriching user experience*. /paper/Inter-Contextual-Distributed-Participatory-Design-Obendorf-Janneck/b1cb3b4d0f3564ed6bf09c7e225b66cf6c434859

Oeberst, A., Cress, U., Back, M., & Nestler, S. (2016). Individual Versus Collaborative Information Processing: The Case of Biases in Wikipedia. In U. Cress, J. Moskaliuk, & H. Jeong (Eds.), *Mass Collaboration and Education* (pp. 165–185). Springer International Publishing. https://doi.org/10.1007/978-3-319-13536-6_9

Oeberst, A., von der Beck, I., Matschke, C., Ihme, T. A., & Cress, U. (2020). Collectively biased representations of the past: Ingroup Bias in Wikipedia articles about intergroup conflicts. *British Journal of Social Psychology*, *59*(4), 791–818. https://doi.org/10.1111/bjso.12356

Olson, J. S., & Olson, G. M. (2013). Collaboration Readiness. In *Working Together Apart: Collaboration over the Internet* (Vol. 6, pp. 43–37). http://www.morganclaypool.com/doi/abs/10.2200/S00542ED1V01Y201310HCI020

Olson, J. S., Wang, D., Olson, G. M., & Zhang, J. (2017). How People Write Together Now: Beginning the Investigation with Advanced Undergraduates in a Project Course. *ACM Trans. Comput.-Hum. Interact.*, *24*(1), 1–40.

Oostveen, A.-M., & van den Besselaar, P. (2004). From Small Scale to Large Scale User Participation: A Case Study of Participatory Design in E-Government Systems. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.1432508

Ospina-Pinillos, L., Davenport, T. A., Ricci, C. S., Milton, A. C., Scott, E. M., & Hickie, I. B. (2018). Developing a Mental Health eClinic to Improve Access to and Quality of Mental Health Care for Young People: Using Participatory Design as Research Methodologies. *Journal of Medical Internet Research*, *20*(5), e9716. https://doi.org/10.2196/jmir.9716

Ospina-Pinillos, L., Davenport, T., Diaz, A. M., Navarro-Mancilla, A., Scott, E. M., & Hickie, I. B. (2019). Using Participatory Design Methodologies to Co-Design and Culturally Adapt the Spanish Version of the Mental Health eClinic: Qualitative Study. *Journal of Medical Internet Research*, *21*(8), e14127. https://doi.org/10.2196/14127

Pace, A., Baligand, L., Haklev, S., Olsen, J. K., de Grez, N., & De Wever, B. (2018). Quantifying Collaboration in Synchronous Document Editing. In *Proceedings of the 3rd Swiss Text Analytics Conference (Swiss-Text 2018)* (CONF). Swiss-Text. https://infoscience.epfl.ch/record/255287

Panciera, K., Halfaker, A., & Terveen, L. (2009). Wikipedians are born, not made: A study of power editors on Wikipedia. *Proceedings of the 2009 ACM International Conference on Supporting Group Work*, 51–60. https://doi.org/10.1145/1531674.1531682

Parser 2011/Real-time collaboration—MediaWiki. (2022, December 2). https://web.archive.org/web/20221202220841/https://www.mediawiki.org/wiki/Parser_2011/Real-time_collaboration#History

Perez, S. (2020). *Facebook tests Wikipedia-powered information panels, similar to Google, in its search results | TechCrunch*. https://web.archive.org/web/20221119001223/https://techcrunch.com/2020/06/11/facebook-tests-wikipedia-powered-information-panels-similar-to-google-in-its-search-results/

Petrucco, C. (2010). Wikipedia as Training Resource for Developing Digital Competences. *International Journal of Digital Literacy and Digital Competence (IJDLDC)*, *1*(3), 29–35. https://doi.org/10.4018/jdldc.2010070103

Pfeil, U., Zaphiris, P., & Ang, C. S. (2006). Cultural Differences in Collaborative Authoring of Wikipedia. *J. Comput. Mediat. Commun.*, *12*(1), 88–113.

Pickard, A. J. (Ed.). (2013a). Data analysis and research presentation. In *Research Methods in Information* (pp. 263–266). Facet. https://doi.org/10.29085/9781783300235.028

Pickard, A. J. (Ed.). (2013b). Ethnography. In *Research Methods in Information* (pp. 135–148). Facet. https://doi.org/10.29085/9781783300235.016

Pickard, A. J. (Ed.). (2013c). Major research paradigms. In *Research Methods in Information* (pp. 5–24). Facet. https://doi.org/10.29085/9781783300235.004

Proposal:Collaborative Google Wave-based Wikipedia editing—Strategic Planning. (2009). https://web.archive.org/web/20230304095210/https://strategy.wikimedia.org/wiki/Proposal:Collaborative_Google_Wave-based_Wikipedia_editing

Proposal:Etherpad-based editing—Strategic Planning. (2010). https://web.archive.org/web/20220817004942/https://strategy.wikimedia.org/wiki/Proposal:Etherpad-based_editing

Proposal:Improve software, keep up with the times—Strategic Planning. (2009). https://web.archive.org/web/20221204211216/https://strategy.wikimedia.org/wiki/Proposal:Improve_software,_keep_up_with_the_times

Prototyping a Collaborative Wikipedia Training and Editing Tool—Social Ideation and Voting | Tricider. (2022). https://web.archive.org/web/20230227195305/https://www.tricider.com/brainstorming/2e9DBXAPDYV

Randall, D., Harper, R., & Rouncefield, M. (2007). *Fieldwork for Design*. Springer London. https://doi.org/10.1007/978-1-84628-768-8

Randall, D., Rouncefield, M., & Hughes, J. A. (1995). Chalk and cheese: BPR and ethnomethodologically informed ethnography in CSCW. *Proceedings of the Fourth Conference on European Conference on Computer-Supported Cooperative Work*, 325–340.

Raymond, E. (1999). The cathedral and the bazaar. *Knowledge, Technology & Policy*, *12*(3), 23–49. https://doi.org/10.1007/s12130-999-1026-0

Reboot. (2017). *New Editor Experiences—MediaWiki—Summary of Findings from Korean and Czech Wikipedia* [Technical Report.]. Wikimedia Foundation. https://upload.wikimedia.org/wikipedia/commons/0/08/New_Editor_Experiences_summary_of_findings%2C_August_2017.pdf

Redi, M., Johnson, I., Gerlach, M., & Zia, L. (2022). *Address Knowledge Gaps, 3 Years On*. https://doi.org/10.6084/m9.figshare.19589662.v3

Rezgui, A., & Crowston, K. (2018). Stigmergic Coordination in Wikipedia. *Proceedings of the 14th International Symposium on Open Collaboration*, *Article 19*, 1–12. https://doi.org/10.1145/3233391.3233543

Rijshouwer, E., Uitermark, J., & de Koster, W. (2021). Wikipedia: A self-organizing bureaucracy. *Information, Communication & Society*, *0*(0), 1–18. https://doi.org/10.1080/1369118X.2021.1994633

Rodero, L. G. (2017). Collaborative work with wikis: Analysis of some innovative educational centers. *Proceedings of the 5th International Conference on Technological Ecosystems for Enhancing Multiculturality*, 1–9. https://doi.org/10.1145/3144826.3145419

Rouse, W. B., Cannon-Bowers, J. A., & Salas, E. (1992). The role of mental models in team performance in complex systems. *IEEE Transactions on Systems, Man, and Cybernetics*, *22*(6), 1296–1308. https://doi.org/10.1109/21.199457

Rubart, J. (2017). Synchronous Collaborative Text Editing in Wikis. *HT (Extended Proceedings)*.

Ruprechter, T., Horta Ribeiro, M., Santos, T., Lemmerich, F., Strohmaier, M., West, R., & Helic, D. (2021). Volunteer contributions to Wikipedia increased during COVID-19 mobility restrictions. *Scientific Reports*, *11*(1), Article 1. https://doi.org/10.1038/s41598-021-00789-3

Said, M. N. H. M., Forret, M., & Eames, C. (2014). Analysis of Contradictions in Online Collaborative Learning using Activity Theory as Analytical Framework. *Jurnal Teknologi*, *68*(2), Article 2. https://doi.org/10.11113/jt.v68.2910

Salvador, T., & Mateas, M. (1997). Introduction to design ethnography. *CHI '97 Extended Abstracts on Human Factors in Computing Systems*, 166–167. https://doi.org/10.1145/1120212.1120325

Sanders, E. B.-N., Brandt, E., & Binder, T. (2010). A framework for organizing the tools and techniques of participatory design. *Proceedings of the 11th Biennial Participatory Design Conference*, 195–198. https://doi.org/10.1145/1900441.1900476

Shane-Simpson, C., & Gillespie-Lynch, K. (2017). Examining potential mechanisms underlying the Wikipedia gender gap through a collaborative editing task. *Computers in Human Behavior*, *66*, 312–328. https://doi.org/10.1016/j.chb.2016.09.043

Shapiro, D. (1994). The limits of ethnography: Combining social sciences for CSCW. *Proceedings of the 1994 ACM Conference on Computer Supported Cooperative Work*, 417–428. https://doi.org/10.1145/192844.193064

Shapiro, D. (2005). Participatory design: The will to succeed. *Proceedings of the 4th Decennial Conference on Critical Computing: Between Sense and Sensibility*, 29–38. https://doi.org/10.1145/1094562.1094567

Shirky, C. (2008). *Here comes everybody: The power of organizing without organizations*. Penguin Press.

Singhal, A. (2012). Introducing the Knowledge Graph: Things, not strings. In *Google*. https://blog.google/products/search/introducing-knowledge-graph-things-not/

Slattery, S. P. (2009). ‘edit this page’: The socio-technological infrastructure of a wikipedia article. *Proceedings of the 27th ACM International Conference on Design of Communication*, 289–296. https://doi.org/10.1145/1621995.1622052

Sliger Krause, R., Rosenzweig, J., & Victor, P. (2017). Out of the Vault: Developing a Wikipedia Edit-a-thon to Enhance Public Programming for University Archives and Special Collections. *Journal of Western Archives, 8*(1). https://digitalcommons.usu.edu/westernarchives/vol8/iss1/3

Slingerland, G., Murray, M., Lukosch, S., McCarthy, J., & Brazier, F. (2022). Participatory Design Going Digital: Challenges and Opportunities for Distributed Place-Making. *Computer Supported Cooperative Work (CSCW)*. https://doi.org/10.1007/s10606-022-09438-3

Statista. (2021). *Global top websites by unique visits 2021 | Statista*. https://web.archive.org/web/20230120101727/https://www.statista.com/statistics/1201889/most-visited-websites-worldwide-unique-visits/

Strategy/Wikimedia movement/2017—Meta. (2017). https://web.archive.org/web/20230306013440/https://meta.wikimedia.org/wiki/Strategy/Wikimedia_movement/2017

Strategy/Wikimedia movement/2018-20/Recommendations/Ensure Equity in Decision-making—Meta. (2018). https://web.archive.org/web/20220729103457/https://meta.wikimedia.org/wiki/Strategy/Wikimedia_movement/2018-20/Recommendations/Ensure_Equity_in_Decision-making

Strategy/Wikimedia movement/2018-20/Recommendations/Improve User Experience—Meta. (2018). https://web.archive.org/web/20211209164426/https://meta.wikimedia.org/wiki/Strategy/Wikimedia_movement/2018-20/Recommendations/Improve_User_Experience

Strategy/Wikimedia movement/2018-20/Recommendations/Improve User Experience—Meta. (2022, September 22). https://web.archive.org/web/20220922015105/https://meta.wikimedia.org/wiki/Strategy/Wikimedia_movement/2018-20/Recommendations/Improve_User_Experience

Structured Discussions—MediaWiki. (2022, November 27). https://web.archive.org/web/20221127101138/https://www.mediawiki.org/wiki/Structured_Discussions#History

Suh, B., Convertino, G., Chi, E. H., & Pirolli, P. (2009). The singularity is not near: Slowing growth of Wikipedia. *Proceedings of the 5th International Symposium on Wikis and Open Collaboration*, 1–10. https://doi.org/10.1145/1641309.1641322

Sullivan, D. (2020). *Google's Knowledge Graph and Knowledge Panels | Google*. https://web.archive.org/web/20221222223332/https://www.blog.google/products/search/about-knowledge-graph-and-knowledge-panels/

Syberthon 2020/Edit-a-thon—Meta. (2020). https://web.archive.org/web/20220626000331/https://meta.wikimedia.org/wiki/Syberthon_2020/Edit-a-thon#Etherpad

Sydow, M., Baraniak, K., & Teisseyre, P. (2017). Diversity of editors and teams versus quality of cooperative work: Experiments on wikipedia. *Journal of Intelligent Information Systems*, *48*(3), 601–632. https://doi.org/10.1007/s10844-016-0428-1

Talk pages consultation 2019—MediaWiki. (2019). https://web.archive.org/web/20230218173709/https://www.mediawiki.org/wiki/Talk_pages_consultation_2019

Talk pages consultation 2019—MediaWiki. (2019). https://web.archive.org/web/20230218173709/https://www.mediawiki.org/wiki/Talk_pages_consultation_2019

Talk:Reading/Web/Desktop Improvements—MediaWiki. (2023). https://web.archive.org/web/20230226181527/https://www.mediawiki.org/w/index.php?title=Talk:Reading/Web/Desktop_Improvements&oldid=5737142

Team Practices Group/Design engagement/Post-interview roadmap—MediaWiki. (2022, December 2). https://web.archive.org/web/20221202222857/https://www.mediawiki.org/wiki/Team_Practices_Group/Design_engagement/Post-interview_roadmap

TeBlunthuis, N., Shaw, A., & Hill, B. M. (2018). Revisiting 'The Rise and Decline' in a Population of Peer Production Projects. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 1–7. https://doi.org/10.1145/3173574.3173929

Teli, M., Foth, M., Sciannamblo, M., Anastasiu, I., & Lyle, P. (2020). Tales of Institutioning and Commoning: Participatory Design Processes with a Strategic and Tactical Perspective. *Proceedings of the 16th Participatory Design Conference 2020 - Participation(s) Otherwise - Volume 1*, 159–171. https://doi.org/10.1145/3385010.3385020

Thomas, S. (2021, August 3). *Online training for online trainers – a response to COVID-19 | the Wikimedia UK blog!* https://web.archive.org/web/20210803012912/https://blog.wikimedia.org.uk/2020/09/online-training-for-online-trainers-a-response-to-covid-19/

Tinati, R., & Luczak-Roesch, M. (2017). Wikipedia: A complex social machine by Ramine Tinati and Markus Luczak-Roesch with Martin Vesely as coordinator. *SIGWEB Newsletter, Winter*, 1–10.

Toniolo, A., Kastiel, K., & Bebchuk, L. (2023, January 24). How Twitter Pushed Stakeholders Under The Bus. *The Harvard Law School Forum on Corporate Governance*. https://corpgov.law.harvard.edu/2023/01/24/how-twitter-pushed-stakeholders-under-the-bus/

Train the Trainer 2022—Call for participants. (2021, November 25). WMUK. https://wikimedia.org.uk/2021/11/train-the-trainer-2022-call-for-participants/

Turner, P., & Turner, S. (2001). A web of contradictions. *Interacting with Computers*, *14*(1), 1–14. https://doi.org/10.1016/S0953-5438(01)00039-X

Uden, L. (2007). Activity theory for designing mobile learning. *International Journal of Mobile Learning and Organisation*, *1*(1), 81–102. https://doi.org/10.1504/IJMLO.2007.011190

UNESCO Wiki4Women Edit-a-thon (India)—Meta. (2020). https://web.archive.org/web/20230214144126/https://meta.wikimedia.org/wiki/UNESCO_Wiki4Women_Edit-a-thon_%28India%29#Etherpad

Usov, S., Safonov, M., Akbilek, E., & Sorokina, L. (2020). Swot Analysis of Moodle Platform Application in the Assessment of Foreign Language Knowledge. *2020 The 4th International Conference on Education and Multimedia Technology*, 31–34. https://doi.org/10.1145/3416797.3416835

Valdez, R. S., Gibbons, M. C., Siegel, E. R., Kukafka, R., & Brennan, P. F. (2012). Designing consumer health IT to enhance usability among different racial and ethnic groups within

the United States. *Health and Technology*, *2*(4), 225–233. https://doi.org/10.1007/s12553-012-0031-6

Viegas, F. B., Wattenberg, M., Kriss, J., & Ham, F. V. (2007). Talk Before You Type: Coordination in Wikipedia. *2007 40th Annual Hawaii International Conference on System Sciences (HICSS'07)*, 78–78. https://doi.org/10.1109/HICSS.2007.511

Vincent, N., & Hecht, B. J. (2020). A Deeper Investigation of the Importance of Wikipedia Links to the Success of Search Engines. *ArXiv:2004.10265 [Cs]*, *abs/2004.10265*. https://arxiv.org/abs/2004.10265

Vincent, N., Johnson, I., & Hecht, B. (2018). Examining Wikipedia With a Broader Lens: Quantifying the Value of Wikipedia's Relationships with Other Large-Scale Online Communities. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 1–13. https://doi.org/10.1145/3173574.3174140

VisualEditor/Portal—MediaWiki. (2022). https://web.archive.org/web/20220312000451/https://www.mediawiki.org/wiki/VisualEditor/Portal

VisualEditor/Real-time collaboration/Timeline—MediaWiki. (2022). https://web.archive.org/web/20230223123814/https://www.mediawiki.org/wiki/VisualEditor/Real-time_collaboration/Timeline

Vora, P., Komura, N., & Stanton Usability Team. (2010). The n00b Wikipedia Editing Experience. *Proceedings of the 6th International Symposium on Wikis and Open Collaboration*, 1–3. https://doi.org/10.1145/1832772.1841393

Voss, A., Procter, R., Slack, R., Hartswood, M., & Rouncefield, M. (2009). Design as and for Collaboration: Making Sense of and Supporting Practical Action. In M. Büscher, R. Slack, M. Rouncefield, R. Procter, M. Hartswood, & A. Voss (Eds.), *Configuring User-Designer Relations: Interdisciplinary Perspectives* (pp. 31–58). Springer. https://doi.org/10.1007/978-1-84628-925-5_3

Wadley, G., Lederman, R., Gleeson, J., & Alvarez-Jimenez, M. (2013). Wadley. *Proceedings of the 25th Australian Computer-Human Interaction Conference: Augmentation, Application, Innovation, Collaboration*, 517–526. https://doi.org/10.1145/2541016.2541030

Wagner, C., García, D., Jadidi, M., & Strohmaier, M. (2015). It's a Man's Wikipedia? Assessing Gender Inequality in an Online Encyclopedia. *CoRR, abs/1501.06307*. http://arxiv.org/abs/1501.06307

Wagner, C., Garcia, D., Jadidi, M., & Strohmaier, M. (2015). It's a Man's Wikipedia? Assessing Gender Inequality in an Online Encyclopedia. *Proceedings of the International AAAI Conference on Web and Social Media*, *9*(1). https://ojs.aaai.org/index.php/ICWSM/article/view/14628

Wagner, C., Graells-Garrido, E., & García, D. (2016). Women Through the Glass-Ceiling: Gender Asymmetries in Wikipedia. *EPJ Data Science, abs/1601.04890*(1), 1–24.

Waiting for Real-Time Collaboration. (2017, August 26). [Webm]. YouTube: Wikimania 2017: Waiting for Real-Time Collaboration – View/save archived versions on archive.org and archive.today. https://web.archive.org/web/20230224154348/https://commons.wikimedia.org/wiki/File:Wikimania_2017-_Waiting_for_Real-Time_Collaboration.webm

Walsh, G., Druin, A., Guha, M. L., Bonsignore, E., Foss, E., Yip, J. C., Golub, E., Clegg, T., Brown, Q., Brewer, R., Joshi, A., & Brown, R. (2012). DisCo: A co-design online tool for asynchronous distributed child and adult design partners. *Proceedings of the 11th International Conference on Interaction Design and Children*, 11–19. https://doi.org/10.1145/2307096.2307099

Wang, D., Olson, J. S., Zhang, J., Nguyen, T., & Olson, G. M. (2015). DocuViz: Visualizing Collaborative Writing. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 1865–1874.

Wiki Education. (2022, August 12). Wiki Education Impact. https://web.archive.org/web/20230722175902/https://wikiedu.org/impact/

Wikimedia 2030 Movement Strategy Recommendations. (2022). https://web.archive.org/web/20230127190207/https://meta.wikimedia.org/wiki/Strategy/Wikimedia_movement/2018-20/Recommendations

Wikimedia Foundation. (2011). *Editor Survey Report*. Wikimedia Foundation. https://web.archive.org/web/20220331052822/https://upload.wikimedia.org/wikipedia/commons/7/76/Editor_Survey_Report_-_April_2011.pdf

Wikimedia Foundation Medium-term plan 2019/Platform evolution—Meta. (2019). Wikimedia Foundation Medium-Term Plan 2019 - Platform Evolution. https://web.archive.org/web/20220413152950/https://meta.wikimedia.org/wiki/Wikimedia_Foundation_Medium-term_plan_2019/Platform_evolution

Wikimedia Foundation Website. (2020). https://web.archive.org/web/20220601011343/https://wikimediafoundation.org/

Wikimedia Growth team. (2022). https://web.archive.org/web/20220929120300/https://www.mediawiki.org/wiki/Growth

Wikimedia Statistics—English Wikipedia - Active editors. (2022). Wikipedia. https://stats.wikimedia.org/#/en.wikipedia.org/contributing/active-editors/normal|line|2002-08-10~2023-02-25|(page_type)~content*non-content|monthly

Wikimedia Vision. (2022). Wikimedia Foundation Vision. https://web.archive.org/web/20220611220914/https://wikimediafoundation.org/about/vision/

Wikipedia Gets a Fresh New Look: First Desktop Update in a Decade Puts Usability at the Forefront – Wikimedia Foundation. (2023, January 29). https://web.archive.org/web/20230129221223/https://wikimediafoundation.org/news/2023/01/18/wikipedia-gets-a-fresh-new-look-first-desktop-update-in-a-decade-puts-usability-at-the-forefront/

Wikipedia talk:Flow. (2023, February 7). Wikipedia. https://web.archive.org/web/20221225004930/https://en.wikipedia.org/wiki/Wikipedia_talk:Flow/Archive_index

Wikipedia talk:Flow/Archive 15—Wikipedia. (2016). https://web.archive.org/web/20220704210330/https://en.wikipedia.org/wiki/Wikipedia_talk:Flow/Archive_15#Flow_removed_from_Enwiki_for_now

Wikipedia:About. (2022, July 25). Wikipedia. https://en.wikipedia.org/w/index.php?title=Wikipedia:About&oldid=1100365834

Wikipedia:Administrators—Wikipedia. (2022). Wikipedia. https://web.archive.org/web/20230109104822/https://en.wikipedia.org/wiki/Wikipedia:Administrators

Wikipedia:FAQ/Contributing—Wikipedia. (2021). https://web.archive.org/web/20210514062552/https://en.wikipedia.org/wiki/Wikipedia:FAQ/Contributing

Wikipedia:Flow—Wikipedia. (2022, December 1). https://web.archive.org/web/20221201102427/https://en.wikipedia.org/wiki/Wikipedia:Flow

Wikipedia:Gender bias and editing on Wikipedia—Wikipedia. (2022, June 3). https://web.archive.org/web/20220603072201/https://en.wikipedia.org/wiki/Wikipedia:Gender_bias_and_editing_on_Wikipedia

Wikipedia:Systemic bias—Wikipedia. (2022, December 26). https://web.archive.org/web/20221226033359/http://en.wikipedia.org/wiki/Wikipedia%3ASystemic_bias

Wikipedia:Village pump (WMF)—Wikipedia. (2022). https://web.archive.org/web/20220527065027/https://en.wikipedia.org/wiki/Wikipedia:Village_pump_(WMF)

Wilson, C. (2013). *Brainstorming and Beyond: A User-Centered Design Method* (Illustrated edition). Morgan Kaufmann.

Wilson, J. (2000). Volunteering. *Annual Review of Sociology*, *26*(1), 215–240. https://doi.org/10.1146/annurev.soc.26.1.215

Wittenbaum, G. M., Vaughan, S. I., & Strasser, G. (2002). Coordination in Task-Performing Groups. In R. S. Tindale, L. Heath, J. Edwards, E. J. Posavac, F. B. Bryant, Y. Suarez-Balcazar, E. Henderson-King, & J. Myers (Eds.), *Theory and Research on Small Groups* (pp. 177–204). Springer US.

Wöhner, T., & Peters, R. (2009, January 1). *Assessing the quality of Wikipedia articles with lifecycle based metrics*. https://doi.org/10.1145/1641309.1641333

Yun, J., Lee, S. H., & Jeong, H. (2019). Early onset of structural inequality in the formation of collaborative knowledge in all Wikimedia projects. *Nature Human Behaviour*, *3*(2), 155–163. https://doi.org/10.1038/s41562-018-0488-z

Zou, J., & Schiebinger, L. (2018). AI can be sexist and racist—It's time to make it fair. *Nature*, *559*(7714), Article 7714. https://doi.org/10.1038/d41586-018-05707-8

Zurita, G., & Nussbaum, M. (2007). A conceptual framework based on Activity Theory for mobile CSCL. *British Journal of Educational Technology*, *38*(2), 211–235. https://doi.org/10.1111/j.1467-8535.2006.00580.x

Appendix A Ethical Approval for Research

A.1 Studying Collaborative Editing Processes in Co-Located Settings

University of St Andrews | FOUNDED 1413

University Teaching and Research Ethics Committee

02 February 2021

Dear Abd,

Thank you for submitting your ethical application, which was considered by the School of Computer Science Ethics Committee on Friday 14th June, where the following documents were reviewed:

1. Ethical Application Form
2. Participant Information Sheet
3. Participant Consent Form
4. Questionnaire
5. Advertisement
6. Ethical Statement
7. Discussion Group Questions

The School of Computer Science Ethics Committee has been delegated to act on behalf of the University Teaching and Research Ethics Committee (UTREC) and has granted this application ethical approval. The particulars relating to the approved project are as follows -

Approval Code:	CS14484	**Approved on:**	08.07.19	**Approval Expiry:**	08.07.2024
Project Title:	Studying Collaborative Wikipedia Editing Processes in Co-Located Settings				
Researcher(s):	Abd Alsattar Ardati				
Supervisor(s):	Alex Voss and Uta Hnrichs				

Approval is awarded for five years. Projects which have not commenced within two years of approval must be re-submitted for review by your School Ethics Committee. If you are unable to complete your research within the five year approval period, you are required to write to your School Ethics Committee Convener to request a discretionary extension of no greater than 6 months or to re-apply if directed to do so, and you should inform your School Ethics Committee when your project reaches completion.

If you make any changes to the project outlined in your approved ethical application form, you should inform your supervisor and seek advice on the ethical implications of those changes from the School Ethics Convener who may advise you to complete and submit an ethical amendment form for review.

Any adverse incident which occurs during the course of conducting your research must be reported immediately to the School Ethics Committee who will advise you on the appropriate action to be taken.

Approval is given on the understanding that you conduct your research as outlined in your application and in compliance with UTREC Guidelines and Policies (http://www.st-andrews.ac.uk/utrec/guidelinespolicies/). You are also advised to ensure that you procure and handle your research data within the provisions of the Data Provision Act 1998 and in accordance with any conditions of funding incumbent upon you.

Yours sincerely

School Ethics Committee Administrator

ethics-cs@st-andrews.ac.uk

The University of St Andrews is a charity registered in Scotland: No SC013532

A.2 Studying Collaborative Editing Processes in Co-Located Settings (Amended)

University of St Andrews | FOUNDED 1413

University Teaching and Research Ethics Committee

20 January 2023

Dear Abd,

Thank you for submitting your amendment application which comprised the following documents:

1. Ethical Amendment Application Form
2. Participant Information Sheet
3. Participant Consent Form

The School of Computer Science Ethics Committee is delegated to act on behalf of the University Teaching and Research Ethics Committee (UTREC) and has approved this ethical amendment application. The particulars of this approval are as follows –

Original Approval Code:	CS14484	**Approved on:**	04.09.19
Amendment Approval Date:	31.10.19	**Approval Expiry Date:**	31.10.24
Project Title:	Studying Collaborative Wikipedia Editing Processes in Co-Located Settings		
Researcher(s):	Abd Alsattar Ardati	**Supervisor(s):**	Alex Voss and Uta Hinrichs

Ethical amendment approval does not extend the originally granted approval period of three years, rather it validate the changes you have made to the originally approved ethical application. If you are unable to complete your research within the original five year validation period, you are required to write to your School Ethics Committee Convener to request a discretionary extension of no greater than 6 months or to re-apply if directed to do so, and you should inform your School Ethics Committee when your project reaches completion.

Any serious adverse events or significant change which occurs in connection with this study and/or which may alter its ethical consideration, must be reported immediately to the School Ethics Committee, and an Ethical Amendment Form submitted where appropriate.

Approval is given on the understanding that you adhere to the 'Guidelines for Ethical Research Practice' (http://www.st-andrews.ac.uk/media/UTRECguidelines%20Feb%2008.pdf).

Yours sincerely,

School Ethics Committee Administrator

ethics-cs@st-andrews.ac.uk

A.3 Prototyping a Collaborative Wikipedia Training and Editing Tool

University of St Andrews | FOUNDED 1413

School of Computer Science Ethics Committee

26 October 2020

Dear Abd,

Thank you for submitting your ethical application which was considered at the School Ethics Committee meeting on 7th October 2020.

The School of Computer Science Ethics Committee, acting on behalf of the University Teaching and Research Ethics Committee (UTREC), has approved this application:

Approval Code:	CS15105	**Approved on:**	26.10.20	**Approval Expiry:**	26.10.25
Project Title:	Prototyping a Collaborative Wikipedia Training and Editing Tool				
Researcher(s):	Abd Alsattar Ardati				
Supervisor(s):	Dr Alexander Voss, Dr Uta Hinrichs and Dr Angela Miguel				

The following supporting documents are also acknowledged and approved:

1. Ethical Application
2. Participant Information Sheet
3. Participant Consent Form
4. Advertisement

Approval is awarded for 5 years, see the approval expiry data above.

If your project has not commenced within 2 years of approval, you must submit a new and updated ethical application to your School Ethics Committee.

If you are unable to complete your research by the approval expiry date you must request an extension to the approval period. You can write to your School Ethics Committee who may grant a discretionary extension of up to 6 months. For longer extensions, or for any other changes, you must submit an ethical amendment application.

You must report any serious adverse events, or significant changes not covered by this approval, related to this study immediately to the School Ethics Committee.

Approval is given on the following conditions:

- that you conduct your research in line with:
  - the details provided in your ethical application
  - the University's Principles of Good Research Conduct
  - the conditions of any funding associated with your work
- that you obtain all applicable additional documents (see the 'additional documents' webpage for guidance) before research commences.

You should retain this approval letter with your study paperwork.

Yours sincerely,

Wendy Boyter
Ethics Administrator

School of Computer Science Ethics Committee
Convenor - Dr Juan Ye. School of Computer Science, Jack Cole Building, North Haugh, KY16 9SX
Telephone: 01334 463253 Email: ethics-cs@st-andrews.ac.uk

A.4 Prototyping a Collaborative Wikipedia Training and Editing Tool (Amended)

University of St Andrews | FOUNDED 1413

School of Computer Science Ethics Committee

12 November 2021

Dear Abd,

Thank you for submitting your ethical application which was considered at the School Ethics Committee on Wednesday 3rd November.

The School of Computer Science Ethics Committee, acting on behalf of the University Teaching and Research Ethics Committee (UTREC), has approved this application:

Approval Code:	CS15836	**Approved on:**	12.11.2021	**Approval Expiry:**	12.11.2026
Project Title:	Prototyping a Collaborative Wikipedia Training and Editing Tool				
Researcher(s):	Abd Alsattar Ardati				
Supervisor(s):	Professor Ian Gent, Dr Alex Voss, Dr Uta Hinrichs and Dr Angela Miguel				

The following supporting documents are also acknowledged and approved:

1. Application Form
2. Participant Information Sheet
3. Participant Consent Form

Approval is awarded for 5 years, see the approval expiry data above.

If your project has not commenced within 2 years of approval, you must submit a new and updated ethical application to your School Ethics Committee.

If you are unable to complete your research by the approval expiry date you must request an extension to the approval period. You can write to your School Ethics Committee who may grant a discretionary extension of up to 6 months. For longer extensions, or for any other changes, you must submit an ethical amendment application.

You must report any serious adverse events, or significant changes not covered by this approval, related to this study immediately to the School Ethics Committee.

Approval is given on the following conditions:

- that you conduct your research in line with:
  - the details provided in your ethical application
  - the University's Principles of Good Research Conduct
  - the conditions of any funding associated with your work
- that you obtain all applicable additional documents (see the 'additional documents' webpage for guidance) before research commences.

You should retain this approval letter with your study paperwork.

School of Computer Science Ethics Committee
Dr Juan Ye/Convenor, Jack Cole Building, North Haugh, St Andrews, Fife, KY16 9SX
Telephone: 01334 463252 Email: ethics-cs@st-andrews.ac.uk
The University of St Andrews is a charity registered in Scotland: No SC013532

Yours sincerely,

SEC Administrator

A.5 Studying Collaborative Online Wikipedia Editing Processes

University of St Andrews | FOUNDED 1413

School of Computer Science Ethics Committee

16 April 2021

Dear Abd,

Thank you for submitting your ethical amendment application.

The School of Computer Science Ethics Committee has approved this ethical amendment application:

Original Approval Code:	CS14484	**Original Approval Date:**	08.07.2019
Amendment Approval Date:	16.04.2021	**Approval Expiry Date:**	08.07.2024
Project Title:	Studying Collaborative Online Wikipedia Editing Processes		
Researcher(s):	Abd Alsattar Ardati, Dr Sara Thomas and Dr Richard Nevell	**Supervisor/PI:**	Dr Alex Voss, Dr Uta Hinrichs and Dr Angela Miguel
School/Unit:	Computer Science		

The following supporting documents are also acknowledged and approved:

1. Ethical Amendment Application Form
2. Original Application Form
3. Participant Information Sheet
4. Participant Consent Form
5. Advertisement
6. Questionnaire
7. Discussion Group Questions

This approval does not extend the originally granted approval period. If you require an extension to the approval period, you can write to your School Ethics Committee who may grant a discretionary extension of no greater than 6 months. For longer extensions, or for any further changes, you must submit an additional ethical amendment application. For all extensions, you should inform the School Ethics Committee when your study is complete.

You must report any serious adverse events, or significant changes not covered by this approval, related to this study immediately to the School Ethics Committee.

Approval is given on the following conditions:

- that you conduct your research in line with:
  - the details provided in your ethical amendment application (and the original ethical application where still relevant)
  - the University's Principles of Good Research Conduct
  - the conditions of any funding associated with your work
- that you obtain all applicable additional documents (see the 'additional documents' webpage for guidance) before research commences.

School of Computer Science Ethics Committee
Dr Juan Ye/Convenor, Jack Cole Building, North Haugh, St Andrews, Fife, KY16 9SX
T: 01334 463252 E: ethics-cs@st-andrews.ac.uk
The University of St Andrews is a charity registered in Scotland: No SC013532

You should retain this approval letter with your study paperwork.

Yours sincerely,

SEC Administrator

University of St Andrews | FOUNDED 1413

School of Computer Science Ethics Committee

30 June 2021

Dear Alex,

Thank you for submitting your ethical application which was considered by the School Ethics Committee.

The School of Computer Science Ethics Committee, acting on behalf of the University Teaching and Research Ethics Committee (UTREC), has approved this application:

Approval Code:	CS15614	**Approved on:**	30.06.2021	**Approval Expiry:**	30.06.2026
Project Title:	Wikipedia as a socio-technical infrastructure and questions of equity				
Researcher(s):	Alexander Voss, Abd Alsattar Ardati, Yi Duan, Yunqian Mao and Georgi Koychev				
Supervisor(s):	Loraine Clarke (for Koychev) and Michael Torpey for (Duan and Mao)				

The following supporting documents are also acknowledged and approved:

1. Application Form

Approval is awarded for 5 years, see the approval expiry data above.

If your project has not commenced within 2 years of approval, you must submit a new and updated ethical application to your School Ethics Committee.

If you are unable to complete your research by the approval expiry date you must request an extension to the approval period. You can write to your School Ethics Committee who may grant a discretionary extension of up to 6 months. For longer extensions, or for any other changes, you must submit an ethical amendment application.

You must report any serious adverse events, or significant changes not covered by this approval, related to this study immediately to the School Ethics Committee.

Approval is given on the following conditions:

- that you conduct your research in line with:
  - the details provided in your ethical application
  - the University's Principles of Good Research Conduct
  - the conditions of any funding associated with your work
- that you obtain all applicable additional documents (see the 'additional documents' webpage for guidance) before research commences.

You should retain this approval letter with your study paperwork.

Yours sincerely,

SEC Administrator

School of Computer Science Ethics Committee
Dr Juan Ye/Convenor, Jack Cole Building, North Haugh, St Andrews, Fife, KY16 9SX
Telephone: 01334 463252 Email: ethics-cs@st-andrews.ac.uk

Appendix B Research Context

This appendix covers my participation and membership of the Wikipedia community, which gives rise to the opportunity for this research.

First, it covers the impact of pivoting from studying co-located and hybrid collaboration in Wikipedia to studying collaboration models in online Wikipedia training on my research methodology. It treats the earliest part of the research, which was abandoned due to COVID-19, although it holds important information that influences the restructure of this research question and methodology. It covers events that shaped the design of this research question and context, focusing on the online distributed Wikipedia community. After that, the focus turns to the distributed nature of studying this community.

At the start of this PhD research, the focus was to see how having a higher awareness level in face-to-face collaboration would influence the onboarding experience and new editors' sense of belonging to the community. The research also aimed at exploring how the group would negotiate their tasks, what work phases they would go through, their group dynamics, and personal tools preferences. The first stage resulted in questions, such as: what characterises face-to-face collaborative activities when editing/writing Wikipedia articles? How does existing editing technology (on Wikipedia or in general) influence collaborative dynamics during face-to-face editing sessions? What are the potentials of having a hybrid form of face-to-face and distributed collaboration? In addition, the research investigated the existing technology that the Wikipedia community use internally and externally for collaborating. The goal was to design a hybrid form of face-to-face and online collaboration model that was hypothesised to be a way to increase the retention rate of the contributors.

After nine months, researching face-to-face collaboration had come to a halt due to COVID-19. However, this has not affected the core question of how to engage the community in defining their challenges and designing solutions that fit Wikipedia's socio-technical structure. The nine months of research have revealed the importance of acknowledging the socio-technical gap before developing, proposing, and adopting new technology in the Wikipedia community.

During that period, tensions were observed between Wikimedia and the Wikipedia community after rolling out a new design feature, such as the talk pages from a project called the *Flow Project* (2016). Though the resulting tool was successfully used in many Wikis, it was

“heavily criticised, and neither has gained wide acceptance on many of the largest [Wikipedia language versions and other wikis].” (*Talk Pages Consultations - MediaWiki*, 2019). The project development stopped in 2019, the same year when a “global consultation with [...] Wikimedians and wiki-minded people [was initiated to]define better tools for wiki communication.” (Talk Pages Consultations - MediaWiki, 2019; Structured Discussions - MediaWiki, 2022). This showed the importance of understanding the community’s views by participating as a member.

IT facilitators are encouraged to be acquainted, and sometimes immersed, in an environment where the system is co-designed with other community members. Becoming a member in such an environment is considered “the key attribute that such facilitators must possess is an ability to listen and learn in cooperation with fellow members.” (Hartswood et al., 2008, p. 61).

I have learned editing Wikipedia independently, and my experience was heightened through the ethnographical study of the Wikipedian fan translation community, taking place during my MSc project (Ardati, 2018). However, during the early planning phase of this research, I attended a training workshop organised by Wikimedia UK. This was the first time I met a Wikipedia trainer. Getting trained that day inspired me to learn more about this environment and its ability for bring change to Wikipedia.

Seeking answers on how to enhance the newcomer using new technology, I applied to be an accredited Wikimedia UK trainer. This significantly improved my understanding of the potential that Wikipedia training has in meeting Wikipedia’s community needs. My focus on the training environment was due to two main factors:

- First, improving Wikipedia training brings positive change to Wikipedia and its community by increasing the number of editors, however, with the remaining challenges in retaining them (Farzan et al., 2016).
- Second, since the training focuses on introducing trainees to *how* the community works,
    - it would be a great medium to observe and learn about the current practices;
    - and offers a suitable medium to present the collaborative writing change aspirations to the trainees who would have the least reluctance to change and pre-conceived ideas on the status quo of editing modes, something quite evident when it comes to experienced editors (Reboot, 2017).

After receiving ethical approval for studying the collaborative Wikipedia editing process in co-located settings in 2019 and becoming an accredited Wikimedia trainer, I ran two training events. First, I ran a pilot event for three PhD students who later collaborated on writing an article in the same room. The participants were free to arrange their environment, assign tasks, and choose their tools. Observing the participants has helped me identify a behaviour that affected my decisions and research direction. For example, the event started with the group rearranging their table to sit next to each other and having their laptops aligned next to each other in a row, so they could view each other's screens. Then, it progressed to choosing the roles of each one and ended with writing an article collaboratively about Agnes Blackadder, a Scottish medical doctor and the first female graduate at the University of St Andrews ('Agnes Forbes Blackadder', 2022). The university is famed for being Scotland's first university, yet providing women with studying opportunities happened only 100 years ago (mbmusablog, 2018). Before creating the article, Wikipedia did have an article on Agnes Blackadder Hall (the student Hall named to Honour the notable student),[131] but none on Agnes Blackadder (*Agnes Blackadder Hall - Wikipedia*, 2018). It was ironic and frustrating to see this article missing, which moved me to research the roots of this gap and led me to research the current collaboration model effect on Wikipedia's knowledge gaps, which are discussed in Section 2.1 and throughout this research. Additionally, changing the physical layout reinforced the importance of a high level of awareness and real-time coordination and collaboration in writing Wikipedia articles. Yet, there was much to uncover, especially regarding role allocation and collaboration in writing the articles.

During that period, I had another face-to-face pilot training for three non-academic community members from the local community. I had to move in between participants to provide real-time support on issues they faced with editing. This also made me wonder about the benefits of real-time collaboration in making the trainers' review process easier.

After discussing with Wikimedia UK the development of a simple online tool I built to help ease the training transition online, I received an invitation to collaborate with other trainers on listing tips and designing an online training model to help trainers identify tools and methods that would replace face-to-face training. As a result, a series of recommendations and online training activities were devised that make the most of available online delivery opportunities. During the first stage of the pandemic, I ran one more online training session with an early prototype of a real-

[131] https://en.wikipedia.org/wiki/Agnes_Blackadder_Hall#Naming

time collaborative writing space using the TogetherJS Library[132] and the Jitsi Conference Tool.[133] I later received feedback from the Wikimedia UK members to improve the collaborative editing space. I did not know where previous engagements and primitive prototypes would lead me, however, they have enriched my knowledge about the research domain, bringing me to the research question discussed in Chapter 1, as well as the methods discussed in Section 3.3. However, this thesis will not discuss these previous engagements and prototypes to focus on the new direction of this research, which started with inviting the two staff members from Wikimedia UK to collaborate on this PhD research. Together with another Wikipedian and staff member at the University of St Andrews, we discussed the new challenges, leading me to redesign this research to focus on supporting online training and distributed real-time collaboration.

[132] https://togetherjs.com/
[133] https://jitsi.org/

Appendix C Observational Study

C.1 Participant Information Sheet

Prototyping a Collaborative Wikipedia Training and Editing Tool

Participant Information Sheet

Thank you for your interest in our study! My name is Abd Alsattar Ardati. I would like to participate as an observer in this online Wikipedia training event as part of my PhD research at the School of Computer Science at the University of St Andrews. The project is being supervised by Dr Alexander Voss, Dr Uta Hinrichs and Dr Angela Miguel.

What is the study about?

We will study people's remote Wikipedia training and editing experience to prototype and evaluate a tool that can support Wikipedia trainers.

Do I have to take part?

This information sheet has been written to help you decide if you would like to take part in this Wikipedia study. It is up to you and you alone whether or not to take part. If you do decide to take part, you will be free to withdraw at any time without providing a reason. If you decide to withdraw from this study at any point, all data collected from the event that you attended will be deleted (with the exception of data already included in published material).

What would I be required to do?

Nothing is required from you apart from reading this information sheet and completing the consent form afterwards.

Are there any risks associated with taking part?

Participating in this research does not exceed the risks of taking part in the Wikipedia training event itself. The research data we collect will not contain any information that we can use to identify you. We will only be recording designations such as 'Trainer A, Trainer B' and 'Trainee A, Trainee B'. No personal information or edits, content or details of the training events will be recorded as part of the observational field notes.

Informed consent

If you are a trainner or a trainee at the event, it is important that you are able to give your informed consent before taking part in the project and you will have the opportunity to ask any questions (contact details below) in relation to the research before you provide your consent.

What information about me or recordings of me ('my data') will you be collecting?

PIS_[12/09/2020]_[2.0]_[Collaborative Wikipedia Training]

During the training event, a researcher will be attending as an observer to take anonymous written observational field notes of participants' individual and collaborative editing and training strategies. This will not include your Wikipedia edits (fieldnotes will be collected using word editing software and/or paper). These notes will contain labels to distinguish different speakers, but these are not linked to anything in the research data that could identify an individual.

How will my data be stored, who will have access to it?

Your data will be recorded in an ANONYMOUS form, which means that no-one, including the researchers, could use any reasonably available means to identify you from the data.

The data that will be collected during this study will be accessible by the researchers and the supervisors (listed below). All the collected data will be transferred to a password-protected folder on the University's OneDrive. The consent information will be stored separately in an encrypted folder that will be stored on the storage servers of the School of Computer Science as well as on the researchers' computers.

How will my data be used, and in what form will it be shared further?

Results from this study will be published in my PhD thesis as well as conference proceedings and/or journal articles. Data collected during this study is anonymized.

If shared (published and/or placed in a database accessible by others), your data will be in an ANONYMISED form, which means that no-one could use any reasonably available means to identify you from the data.

The results will be finalised by the end of December 2023 and written up as part of my PhD research.

When will my data be destroyed?

Data is anonymous, and it will be kept indefinitely.

Will my participation be confidential?

We do not collect any identifying information in our research data, so this data will not reflect your participation.

We will not share with third parties any personally identifying information about you. However, other participants of the Wikipedia training event, of course, will know about your participation.

Lawful basis for making use of personal data and data protection rights

The lawful basis that the University will rely on to make use of your personal data during the research and for related research projects in the future, as described to you, is public task; where

special category personal data are used the lawful basis is archiving purposes in the public interest, scientific or historical research purposes or statistical purposes.

The University of St Andrews is a Data Controller for the information you provide about you. You have a range of rights under the data protection legislation, including the right of complaint. However, some of those rights may not be available where you provide personal data for research purposes. For questions, comments or requests, consult the University website at https://www.st-andrews.ac.uk/terms/data-protection/rights/, or email dataprot@st-andrews.ac.uk.

You will be able to withdraw your data before the end of December 2023. Your data is anonymised, if you decided to withdraw, the data from the whole event will be deleted because we will no longer know which data is yours.

Ethical Approvals

This research proposal has been scrutinised and subsequently granted ethical approval by the University of St Andrews Teaching and Research Ethics Committee. The approval reference number is CS15105.

What should I do if I have concerns about this study?

In the first instance and before you move to the consent page, you are encouraged to raise your concerns or questions with the researcher, and if you do not feel comfortable doing so, you should contact the research supervisors. A full outline of the procedures governed by the University Teaching and Research Ethics Committee is available at http://www.st-andrews.ac.uk/utrec/guidelinespolicies/complaints

Contact details

Researcher(s):	Abd Alsattar Ardati aaa8@st-andrews.ac.uk	**Supervisor(s):**	Dr Alexander Voss alex.voss@st-andrews.ac.uk
	Dr Sara Thomas sara.thomas@wikimedia.org.uk		Dr Uta Hinrichs uh3@st-andrews.ac.uk
	Dr Richard Nevell richard.nevell@wikimedia.org.uk		Dr Angela Miguel arm14@st-andrews.ac.uk

C.2 Consent Form

Prototyping a Collaborative Wikipedia Training and Editing Tool

Consent Form

The University of St Andrews attaches high priority to the ethical conduct of research. We therefore ask you to consider the following points before signing this form. Electronically ticking the consent checkbox confirms that you are willing to participate in this study, however, this does not commit you to anything you do not wish to do and you are free to withdraw your participation at any time without giving reasons.

Please initial box

- I understand the contents of the Participant Information Sheet (marked 'PIS_[12-09-2020]_[Collaborative Wikipedia Training]') ☐
- I have been given the opportunity to ask questions about the study and have had them answered satisfactorily. ☐
- I understand that my participation is entirely voluntary and that I can withdraw from the study at any time without giving an explanation. ☐
- I understand who will have access to my data, how it will be stored, in what form it will be shared, and what will happen to it at the end of the study. I understand that If you decide to withdraw from this study at any point, all data collected from the event that I attended will be deleted. ☐
- I understand that part of this research involves taking notes of the training event. The data will be kept securely and stored on the university's OneDrive separately to any identifiable information, i.e. consent forms and my name or Wikipedia username. ☐
- I agree to the anonymized form of the notes to be published as part of this research. ☐
- I give permission for notes to be used in future studies in anonymized form without further consultation. ☐

Signatures			
I confirm that I am willing to take part in this research			
	Print name	Date	**Signature**
Participant			

Studying Collaborative Wikipedia Editing Processes in Co-Located Settings

C.3 Online Qualtrics Survey

Participant Information Sheet

Thank you for your interest in our study! My name is Abd Alsattar Ardati and I would like to be observing this online Wikipedia training event as part of my PhD research at the School of Computer Science at the University of St Andrews. The project is being supervised by Dr Alexander Voss, Dr Uta Hinrichs and Dr Angela Miguel.

What is the study about?
We will study people's remote Wikipedia training and editing experience to prototype and evaluate a tool that can support Wikipedia trainers.

Do I have to take part?

This information sheet has been written to help you decide if you would like to take part in this Wikipedia study. It is up to you and you alone whether or not to take part. If you do decide to take part, you will be free to withdraw at any time without providing a reason. If you decide to withdraw from this study at any point, all data collected from the event that you attended will be deleted (with the exception of data already included in published material).

What would I be required to do?

Nothing is required from you apart from reading this information sheet and completing the consent form afterwards.

Are there any risks associated with taking part?

Participating in this research does not exceed the risks of taking part in the Wikipedia training event itself. The research data we collect will not contain any information that we can use to identify you. We will only be recording designations such as 'Trainer A, Trainer B' and 'Trainee A, Trainee B'. No personal information or edits content or details of the training events will be recorded as part of the observational fieldnotes.

Informed consent

If you are a trainner or a trainee at the event, it is important that you are able to give your informed consent before taking part in the project and you will have the opportunity to

ask any questions (contact details below) in relation to the research before you provide your consent.

What information about me or recordings of me ('my data') will you be collecting?

During the training event, a researcher will be attending as an observer to take anonymous written observational fieldnotes of participants' individual and collaborative editing and training strategies this will not inlcude your wikipedia edits (fieldnotes will be collected using word editing software and/or paper). These notes will contain lables to distinguish different speakers, but these are not linked to anything in the research data that could identify an individual.

How will my data be stored, who will have access to it?

Your data will be recorded in **ANONYMOUS** form, which means that no-one, including the researchers, could use any reasonably available means to identify you from the data.

The data that will be collected during this study will be accessible by the researchers and the supervisors (listed below). All the collected data will be transferred to a password-protected folder on the University's OneDrive. The consent information will be stored separately in an encrypted folder that will be stored on the storage servers of the School of Computer Science as well as on the researchers' computers.

How will my data be used, and in what form will it be shared further?

Results from this study will be published in my PhD thesis as well as conference proceedings and/or journal articles. Data collected during this study is anonymized.

If shared (published and/or placed in a database accessible by others), your data will be in an **ANONYMISED** form, which means that no-one could use any reasonably available means to identify you from the data.

The results will be finalised by end of December 2023 and written up as part of my PhD research.

When will my data be destroyed?

Data is anonymous and it will be kept indefinitely.

Will my participation be confidential?

We do not collect any identifying information in our research data, so this data will not reflect your participation.

We will not share with third parties any personally identifying information about you. However, other participants of the Wikipedia training event of course will know about your participation.

Will my participation be confidential?

We do not collect any identifying information in our research data, so this data will not reflect your participation.

We will not share with third parties any personally identifying information about you. However, other participants of the Wikipedia training event of course will know about your participation. Also, since all editing activities on Wikipedia are tracked and usernames are publicly available, other people outside of this event may be able to find out about your participation.

Lawful basis for making use of personal data and data protection rights

The lawful basis that the University will rely on to make use of your personal data during the research and for related research projects in the future, as described to you is public task; where special category personal data are used the lawful basis is archiving purposes in the public interest, scientific or historical research purposes or statistical purposes.

The University of St Andrews is a Data Controller for the information you provide about you. You have a range of rights under the data protection legislation, including the right of complaint. However, some of those rights may not be available where you provide personal data for research purposes. For questions, comments or requests, consult the University website at https://www.st-andrews.ac.uk/terms/data-protection/rights/, or email dataprot@st-andrews.ac.uk.

You will be able to withdraw your data before the end of December 2023. Your data is anonymised, if you decided to withdraw, the data from the whole event will deleted because we will no longer know which data is yours.

Ethical Approvals

This research proposal has been scrutinised and subsequently granted ethical approval by the University of St Andrews Teaching and Research Ethics Committee. The approval reference number is CS15105.

What should I do if I have concerns about this study?

In the first instance and before you move to the consent page, you are encouraged to raise your concerns or questions with the researcher and if you do not feel comfortable doing so, then you should contact the research supervisors. A full outline of the procedures governed by the University Teaching and Research Ethics Committee is available at http://www.st-andrews.ac.uk/utrec/guidelinespolicies/complaints

Contact details

Researcher(s):	Abd Alsattar Ardati aaa8@st-andrews.ac.uk	**Supervisor(s):**	Dr Alexander Voss alex.voss@st-andrews.ac.uk
	Dr Sara Thomas sara.thomas@wikimedia.org.uk		Dr Uta Hinrichs uh3@st-andrews.ac.uk
	Dr Richard Nevell richard.nevell@wikimedia.org.uk		Dr Angela Miguel arm14@st-andrews.ac.uk

This Participant Information Sheet is available for download through the following link, the same link to this document is included in the advertisement, so you have the opportunity to download it at any time:

https://universityofstandrews907-my.sharepoint.com/:b:/g/personal/aaa8_st-andrews_ac_uk/ETjK9IVZe7JHozlCpYedjU4BzhPr5OJ7sdqS9ljMefaL6Q?e=XaKbhc

Consent Form

The University of St Andrews attaches high priority to the ethical conduct of research. We therefore ask you to consider the following points before signing this form. Electronically ticking the consent checkbox confirms that you are willing to participate in this study, however, this does not commit you to anything you do not wish to do and you are free to withdraw your participation at any time without giving reasons.

Please initial box

- ☐ I understand the contents of the Participant Information Sheet provided in the previous page.
- ☐ I have been given the opportunity to ask questions about the study and have had them answered satisfactorily.
- ☐ I understand that my participation is entirely voluntary and that I can withdraw from the study at any time without giving an explanation.
- ☐ I understand who will have access to my data, how it will be stored, in what form it will be shared, and what will happen to it at the end of the study. I understand that If you decide to withdraw from this study at any point, all data collected from the event that I attended will be deleted.
- ☐ I understand that part of this research involves taking notes of the training event. The data will be kept securely and stored on the university's OneDrive separately to any identifiable information, i.e. consent forms and my name or Wikipedia username.
- ☐ I agree to the anonymized form of the notes to be published as part of this research.
- ☐ I give permission for notes to be used in future studies in anonymized form without further consultation.
- ☐ I confirm that I am over 18 years old of age or older (a requirement by the university ethics)

I confirm that I am willing to take part in this research

☐ If you accept please tick the box, and add your name or your Wikipedia account username in the text field below (This is just so that we can track who has consented. It will not be stored with any identifiable data.)

Survey Powered By Qualtrics

C.4 Advertisement

University of St Andrews

Prototyping a Collaborative Wikipedia Training and Editing Tool

Advertisement

As part of the training, we would like to invite you to participate in a research study that we are conducting to better understand how to support collaborative Wikipedia training and to better accommodate newcomers to Wikipedia in terms of their different learning strategies and levels of expertise.

As part of this study, a researcher will be attending the training event as an observer to take some notes.

Participation in this study is voluntary, and we need all of the participants' consent before the researcher can take part in the event. Using the following link, you will find the **Participant Information Sheet.** If you have any question once you finished reading it, please contact any of the researchers below. At the end of it, you will find the **Consent Form**.

https://standrews.eu.qualtrics.com/jfe/form/SV_1M4ljd3v5JkU6tT

The Participant Information Sheet, which is included in the previous form, is available for download through the following link, so you have the opportunity to download it at any time:

https://universityofstandrews907-my.sharepoint.com/:b:/g/personal/aaa8_st-andrews_ac_uk/ETiK9lVZe7JHozlCpYediU4BzhPr5QJ7sdqS9IiMefaL6Q?e=6kPlsf

Contact Details:

Researchers:
Abd Alsattar Ardati (aaa8@st-andrews.ac.uk)
Dr Sara Thomas (sara.thomas@wikimedia.org.uk)
Dr Richard Nevell (richard.nevell@wikimedia.org.uk)

Supervisors:
Dr Alexander Voss (alex.voss@st-andrews.ac.uk)
Dr Uta Hinrichs (uh3@st-andrews.ac.uk)
Dr Angela Miguel (arm14@st-andrews.ac.uk)

School of Computer Science
University of St Andrews
Jack Cole Building
North Haugh
St Andrews KY16 9SX
+44 1334 46 3262

Appendix D Participatory Design Sessions

D.1 Online Qualtrics Survey

It includes the Participant Information Sheet, Consent form and a questionnaire about previous experience. Please note that the survey has a logical flow, so some of the questions are conditional to specific answers.

Participant Information Sheet

What is the study about?

Thank you for your interest in our Wikipedia editing event! My name is Abd Alsattar Ardati, and this event is part of my PhD research at the School of Computer Science. The project is being supervised by Prof Ian Gent, Dr Alexander Voss, Dr Uta Hinrichs, and Dr Angela Miguel.

We would like to invite you to participate in a study that will help us understand collaborative Wikipedia editing processes as part of online events, in order to better support the collaborative editing of Wikipedia articles and other projects.

This survey will include Participant Information, questions to help us know how familiar and experienced are you with Wikipedia, and the consent form.

Based on your answers, we will contact invite you to a group discussion.

Do I have to take part?

This information sheet has been written to help you decide if you would like to participate in this research. It is up to you, and you alone, whether or not to take part. If you do decide to take part, you will be free to withdraw at any time without providing a reason and at no detriment. If you decide to withdraw from this study at any point, all data collected from you will be deleted (with the exception of data already included in published material).Click to write the question text

What would I be required to do?

In this survey, you will complete a 2-3 min questionnaire to assess your familiarity with Wikipedia. Later, together with other participants, you will be asked to participate in a group discussion conducted via a teleconference system such as Microsoft Teams. If you agree, then the group discussion will be video recorded, including the audio. The discussion will vary in themes. However, it will cover topics such as:

- Identifying the need for a Wikipedia training tool that facilitates collaborative editing.
- Identifying the people who could benefit from the tool.
- What defines a successful outcome for the tool?
- How will the tool be used and in which context?
- How should it look, and what should it include? (This will include sketching)
- Evaluating ideas generated by the previous questions

The session will take 1 to 2 hours.

Are there any risks associated with taking part?

Participating in this research does not exceed the risks of everyday life. We will take all precautions to protect participants' identity and privacy through all stages of the Data Lifecycle. However, we cannot guarantee confidentiality as other participants will know about your participation in the event and see your name in the Microsoft Teams.

Informed consent

It is important that you are able to give your informed online consent using Qualtrics, an online survey system recommended by our University, before taking part in my project, and you will have the opportunity to ask any oral or written questions in relation to the research before you provide your consent. For more information about this, please check the last section of this PIS.

What information about me or recordings of me ('my data') will you be collecting?

- Data about your familiarity and experience with Wikipedia are gathered through this questionnaire.
- Audio and video recordings of the group discussion session with other participants (using Microsoft Teams) about the design of the tool. Recording the collaborative editing process will help us collect details that we might miss during the session about your discussion and therefore enable us to better understand the needs of trainers and volunteer editors engaged in Wikipedia editing. The video will include the faces of participants who are taking part in the study. Group discussion might include discussing the tool requirements, sketching, and prototyping.
- If your session includes using the Miro Virtual collaboration board tool, we will gather data from your contributions to Miro https://miro.com/app/dashboard/
- Data regarding your views of the future development of the tool design iterations will be gathered through group discussions and follow-up questionnaires.

How will my data be stored and who will have access to it?

The data that will be collected during this study will be accessible by the researcher and his supervisors (listed below). All data, including audio and video data, will be anonymized prior to any publication or presentation – this means that the participants' user names and faces in Microsoft Teams' recorded videos will be blurred, and verbal statements will be transcribed so that no one can identify you in audio or video stills or reported participant statement. All the collected data will be transferred to a password-protected folder and kept on the University's OneDrive cloud storage system provided by Microsoft.

How will my data be used, and in what form will it be shared further?

Results from this study will be published in conference proceedings and/or journal articles. Any

data collected during this study will be anonymized or pseudonymised prior to any publication or presentation.

Some of your data will be shared (published and/or placed in a database accessible by others) in a PSEUDONYMISED form, which means that your data will be edited so that you are referred to by a unique reference such as a code number or different name. There will be a 'key' document, which will link your unique reference to your real identity. The key will be kept in encrypted form on a password-protected computer, separate from the pseudonymized data, and only the researchers listed in this application will have access to it and be able to reconnect your data to you at a later date.
The results will be finalised by end of December 2023 and written up as part of my PhD research.

When will my data be destroyed?

Your data will be shared as described above and all raw data collected will be stored for a period of 5 years before being destroyed along with the pseudoanonymised key and videos. Anonymized data will be kept indefinitely.

Will my participation be confidential?

We might include the name of the participant role, but we will not be publishing any names. Instead, we will use designations such as 'Wikipedian A, Wikipedian B'. Any sensitive shared information, which is not in Public knowledge, and may affect the participants will be removed. We will not share with third parties any personal identifying information about you. However, other participants of the group discussion of course will know about your participation, and we cannot guarantee or control what data the third-party tools collect.

Lawful basis for making use of personal data and data protection rights

The lawful basis that the University will rely on to make use of your personal data during the research and for related research projects in the future, as described to you is public task; where special category personal data are used the lawful basis is archiving purposes in the public interest, scientific or historical research purposes or statistical purposes.
The University of St Andrews is a Data Controller for the information you provide about you. You have a range of rights under the data protection legislation, including the right of complaint. However, some of those rights may not be available where you provide personal data for research purposes. For questions, comments or requests, consult the University website at https://www.st-andrews.ac.uk/terms/data-protection/rights/, or email dataprot@st-andrews.ac.uk.
You will be able to withdraw your data before the end of December 2023. If your data is anonymised, we will not be able to withdraw it after that point, because we will no longer know which data is yours. Also, we won't be able to withdraw your data participation if publications have happened before your decision to withdraw.

Ethical Approvals

This research proposal has been scrutinised and subsequently granted ethical approval by the University of St Andrews Teaching and Research Ethics Committee. The approval reference number is CS15836.

What should I do if I have concerns or questions about this study?

In the first instance, you are encouraged to raise your concerns or question by emailing the principal researcher. If you do not feel comfortable doing so, then you should contact my primary supervisor, Prof Ian Gent. A full outline of the procedures governed by the University Teaching and Research Ethics Committee is available at http://www.st-andrews.ac.uk/utrec/guidelinespolicies/complaints

Contact details

Researcher(s):	Abd Alsattar Ardati aaa8@st-andrews.ac.uk	**Supervisor(s):**	Dr Alexander Voss alex.voss@st-andrews.ac.uk
	Dr Sara Thomas sara.thomas@wikimedia.org.uk		Dr Uta Hinrichs uh3@st-andrews.ac.uk
	Dr Richard Nevell richard.nevell@wikimedia.org.uk		Dr Angela Miguel arm14@st-andrews.ac.uk
	Dr Kirsty Ross ksrh1@st-andrews.ac.uk		Prof Ian Gent ian.gent@st-andrews.ac.uk

This Participant Information Sheet and Consent Form are available for download through the following link, the same PIS link is included in the advertisement, so you have the opportunity to download it at any time:

Downloadable PIS link for your reference. Pis 08 07 2021 1.0 collaborative wikipedia training

Consent Form

The University of St Andrews places a high priority on the ethical conduct of research. We therefore ask you to consider the following points before signing this form. Your electronic signature confirms that you are willing to participate in this study, however, signing this form does not commit you to anything you do not wish to do and you are free to withdraw your participation at any time.Click to write the question text

Please initial box

- ☐ I understand the contents of the Participant Information Sheet provided in the previous page.
- ☐ I have been given the opportunity to ask questions about the study and have had them answered satisfactorily.

- [ ] I understand that my participation is entirely voluntary and that I can withdraw from the study at any time without giving an explanation and experience no detriment.
- [ ] I understand who will have access to my data, how it will be stored, in what form it will be shared, and what will happen to it at the end of the study.
- [ ] I understand that I will be able to withdraw data before the end of December 2023 and that, if my data has been anonymised or if publications have happened, it cannot be withdrawn after that point.
- [ ] Part of this research involves associating my ideas from the discussion with identifiers.
- [ ] I confirm that I am over 18 years old of age or older (a requirement of the University Ethics Group).

Video recordings & Wikipedia username

I understand that part of this research involves taking video recordings, including the audio, of the group discussion sessions. These recordings will be kept securely and stored on the University's OneDrive separate from any identifiable information, i.e. consent forms.

Videos can be a valuable resource for future studies, and therefore, we ask for your additional consent to maintain this data for this purpose.

- [ ] I agree to be video recorded as part of this research.
- [ ] I agree to my video material to be published as part of this research in anonymized form.
- [ ] I give consent for video screenshots material to be used in future studies in anonymized form by blurring faces and removing any identifiers without further consultation.

Please type your name or a fake name

I understand that part of this research involves associating my ideas from the discussion with identifiers. This requires researchers to store my email and name or a fake name which will be kept securely and stored separately to any other data collected during this study, i.e., consent forms, questionnaires, or video recordings.

I confirm that I am willing to take part in this research

- [ ] If you accept, please tick the box.

This Consent Form is available for download through the following link:

Downloadable consent form for your reference.

Your previous experience.

Your email address, please.

What gender do you identify as?

- ○ Female
- ○ Male
- ○ Other
- ○ Prefer not to answer

Where in the world are you physically located at the moment (city, country)?

Have you ever edited a Wikipedia article or been involved in any other Wikimedia projects? Please select the answer that is closest to your experience.

- ○ Yes, I have been an active contributor for ____ years. Please enter the number of years below:
- ○ Yes, I have an account, but I have not been active for more than one year.
- ○ Yes, other.
- ○ No, but I have thought about it (please share why you did not edit).
- ○ No, I have never edited or contributed to Wikipedia.
- ○ No, I did not realize that I could.
- ○ No, other.

Have you ever attended Wikipedia training?

- ◯ Yes, I received face to face training.
- ◯ Yes, I received online training.
- ◯ No, I have learned how to edit on my own.
- ◯ No, Other.

How long has it been since you received the training? (e.g. almost two months, a year)

Who provided you with the training? (e.g. Charity, Organisation, or a specific trainer...)

Do you currently have an administrative role in any Wikipedia language version?(or any other Wikimedia projects) before?

- ◯ Yes
- ◯ No, but previously I had.
- ◯ No

What is your role? And in which Wikimedia Project/s?

What was your role? And in which Wikimedia Project/s?

Do you currently have a role in Wikimedia Foundation or its chapters?

- ◯ Yes
- ◯ No, but previously I had.
- ◯ No

What is your role? And in which Wikimedia Chapter/Project?

What was your role? And in which Wikimedia Chapter/Project?

D.2 Advertisement

Advertisement

University of St Andrews

Prototyping a Collaborative Wikipedia Training and Editing Tool

We would like to invite you to participate in a research study we are conducting to better understand how to **support collaborative Wikipedia editing** and to better accommodate newcomers to Wikipedia in terms of their different learning strategies and levels of expertise.

As part of this study, we will be **discussing** the design of an **online collaborative training tool** with groups of participants (18+) from the general public as well as expert Wikipedia editors and trainers.

You are invited to participate in this research and – as part of group work – discuss [topics to be determined]. The event will take place on [date here] at [time here]

Participation in these **1-2 hours sessions** is voluntary. Participants will be compensated with £15 worth of gift vouchers for each session as a thank you for their time [or other compensation options that worth the same value].

During the event, we would like to collect some observational data to support our research. If you are interested in participating, please click on the link below which includes the **Participant Information Sheet**, questions about the person's familiarity and experience with Wikipedia, and the **Electronic Consent Form:**

https://standrews.eu.qualtrics.com/jfe/form/SV_6xqV1dXWoGX5an4

If you have any questions, feel free to contact the researchers. Their details are below:

Contact Details

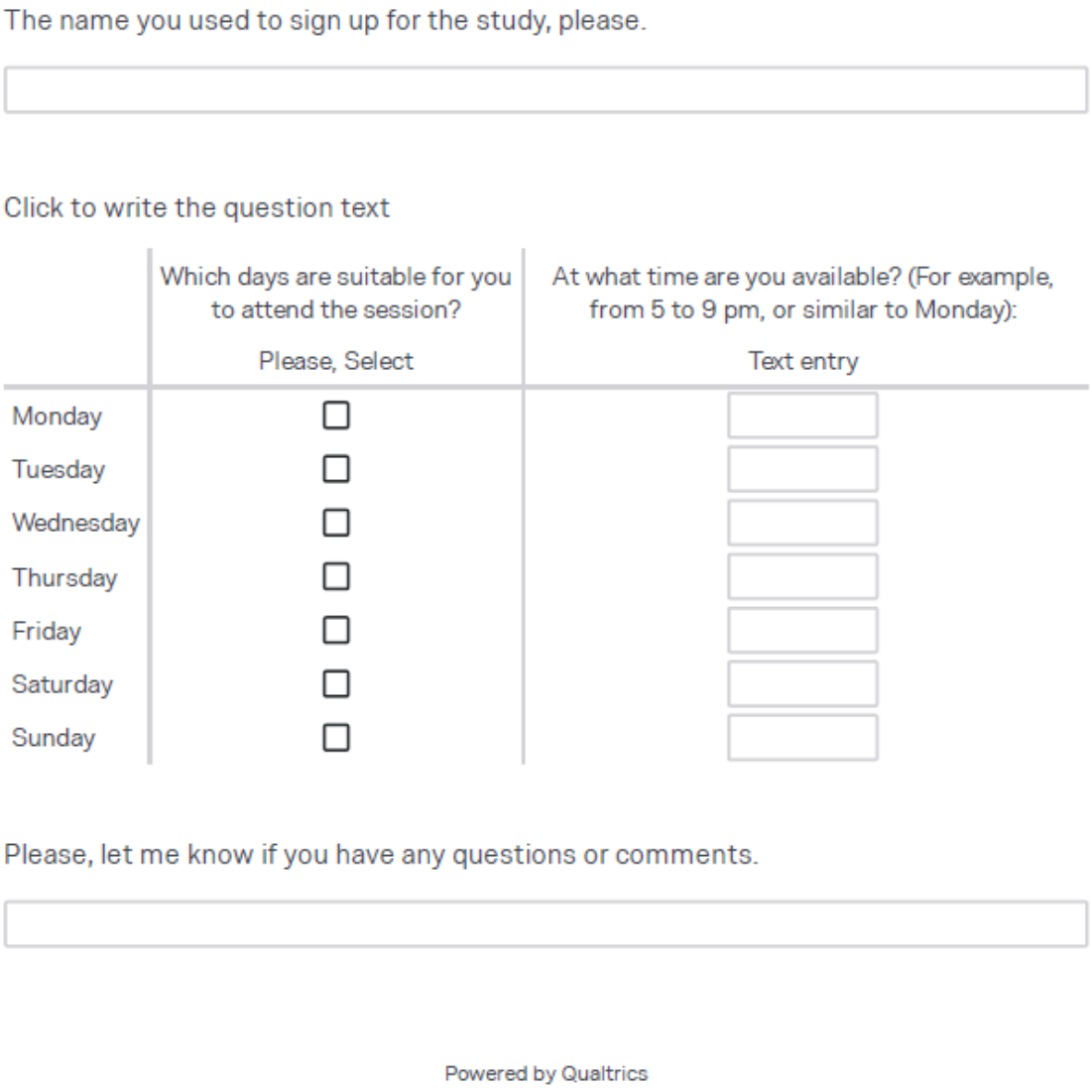
Researchers:
Abd Alsattar Ardati (aaa8@st-andrews.ac.uk)
Dr Sara Thomas (sara.thomas@wikimedia.org.uk)
Dr Richard Nevell (richard.nevell@wikimedia.org.uk)
Dr Kirsty Ross (ksrh1@st-andrews.ac.uk)

Supervisors:
Prof Ian Gent (ian.gent@st-andrews.ac.uk)
Dr Alexander Voss (alex.voss@st-andrews.ac.uk)
Dr Uta Hinrichs (uh3@st-andrews.ac.uk)
Dr Angela Miguel (arm14@st-andrews.ac.uk)

School of Computer Science
University of St Andrews
Jack Cole Building
North Haugh
St Andrews KY16 9SX

+44 1334 46 3262

D.3 Polling for the Session Time and Date Using Qualtrics



The name you used to sign up for the study, please.

Click to write the question text

	Which days are suitable for you to attend the session? Please, Select	At what time are you available? (For example, from 5 to 9 pm, or similar to Monday): Text entry
Monday	☐	
Tuesday	☐	
Wednesday	☐	
Thursday	☐	
Friday	☐	
Saturday	☐	
Sunday	☐	

Please, let me know if you have any questions or comments.

Powered by Qualtrics

D.4 Reflections on Conducting Virtual Sessions

The first session was conducted using Microsoft Office and Miro, a virtual collaboration board. One of the deliverables of this research will be the recommendations on how to set up a session using online tools. Throughout the sessions, there were many takeaways about the research tools used in the first session.

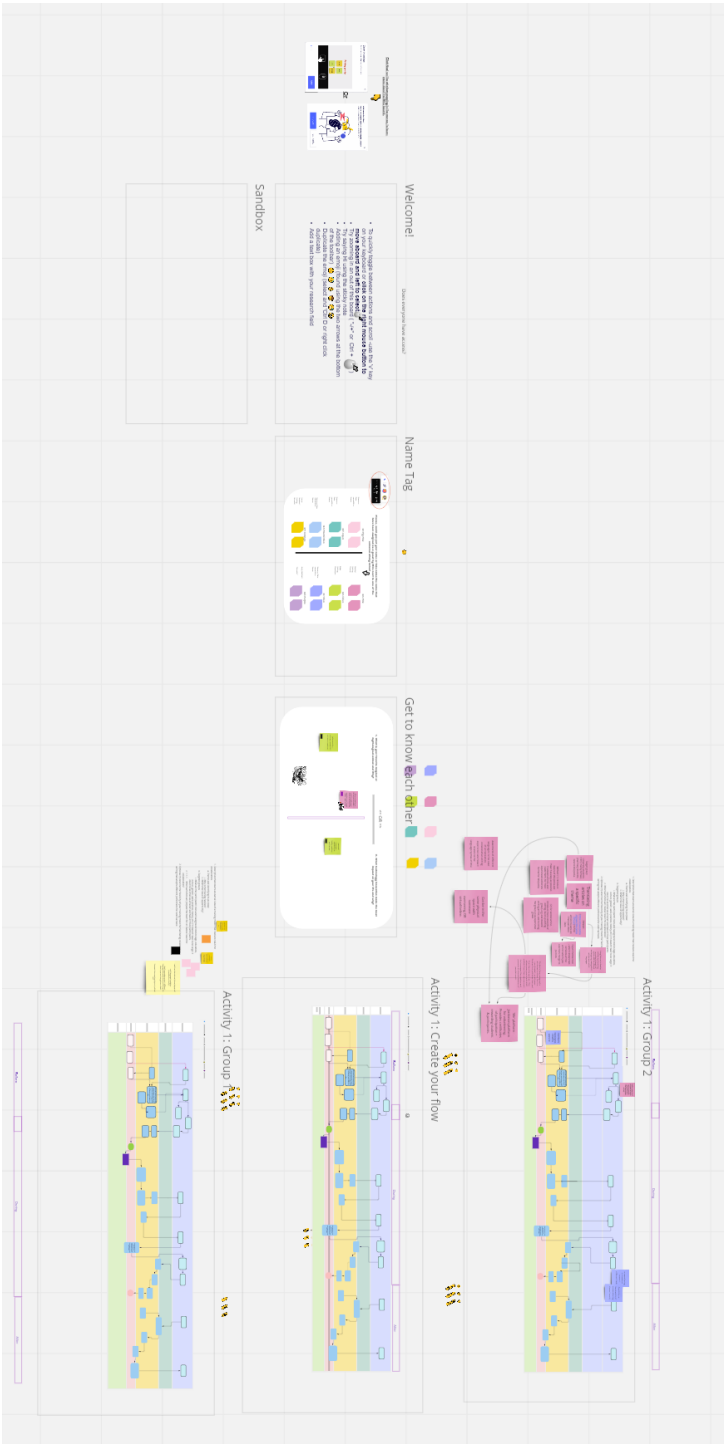
- **Devices:** The sessions were designed assuming that participants had access to more than one screen, having worked with multiple screens for years; thus, the transition between the tools

for participants who had one screen was not smooth. Also, some participants who had small screens did not have the ability to zoom in out, which led them to enlarge some elements (see Figure 5.11, where the size of sticky notes varies).

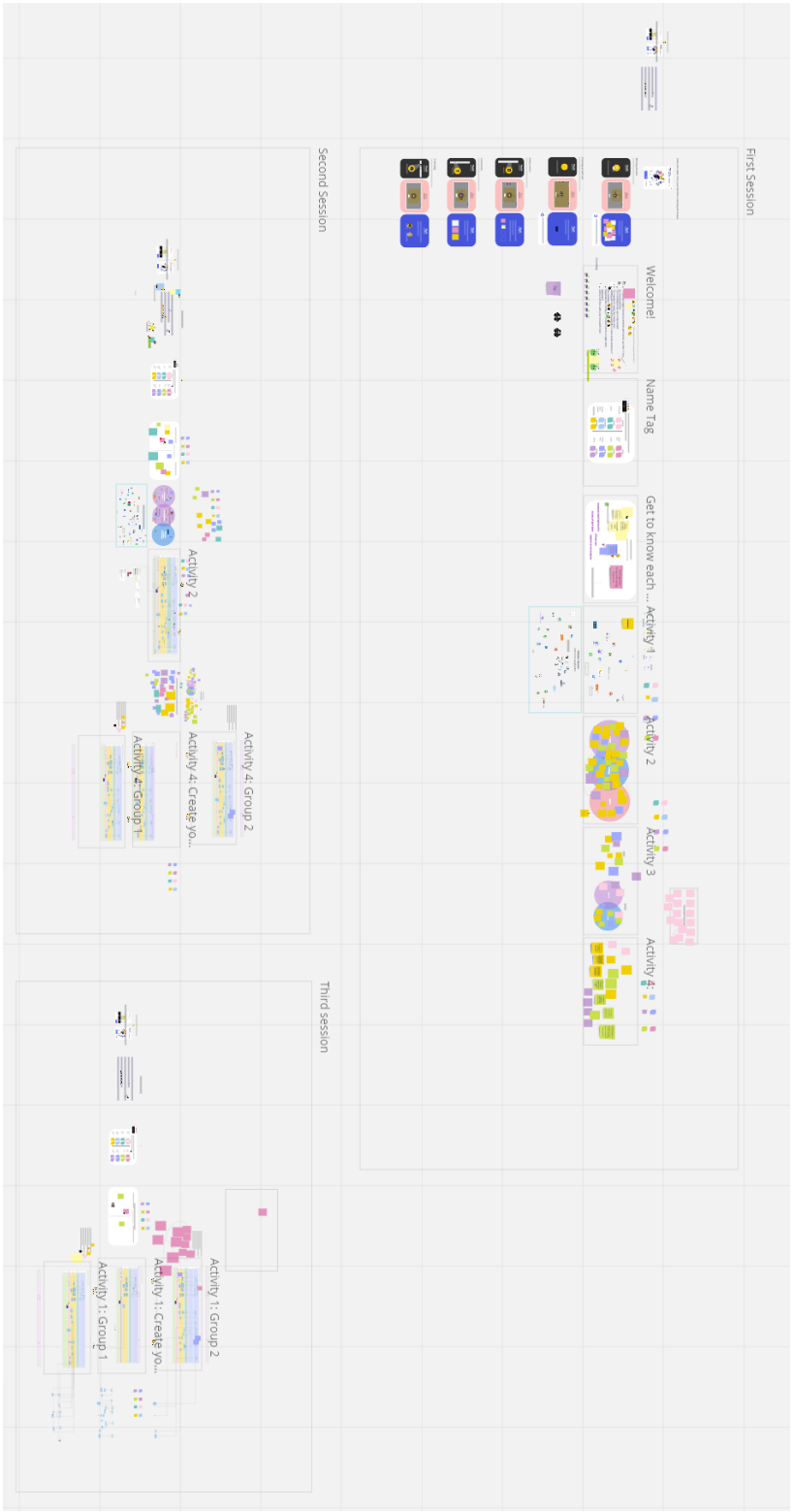
- **Microsoft Teams:** some participants had access to an account different from the one they signed up with, leading to confusion and wasting time at the start. Some participants forgot to mute when not speaking.
- **Miro:**
    - Paying attention to the elements and layers in Miro is important. Since this I was unaware of its "hide and reveal" feature,[134] the activities had to be copied during the session, so I did not want to confuse the participants with all activities at once. Copying on the spot led to some elements covering other layers.
    - From the first onboarding process, people on the Miro board took longer than expected to get accustomed to the environment.
    - Focusing on the Miro board and having too many moving elements and unexpected issues divided my attention and led to losing the opportunities to pick up on the participants' body language in the Microsoft Teams.

[134] https://community.miro.com/product-news-31/introducing-the-new-hide-and-reveal-setting-for-frames-2005

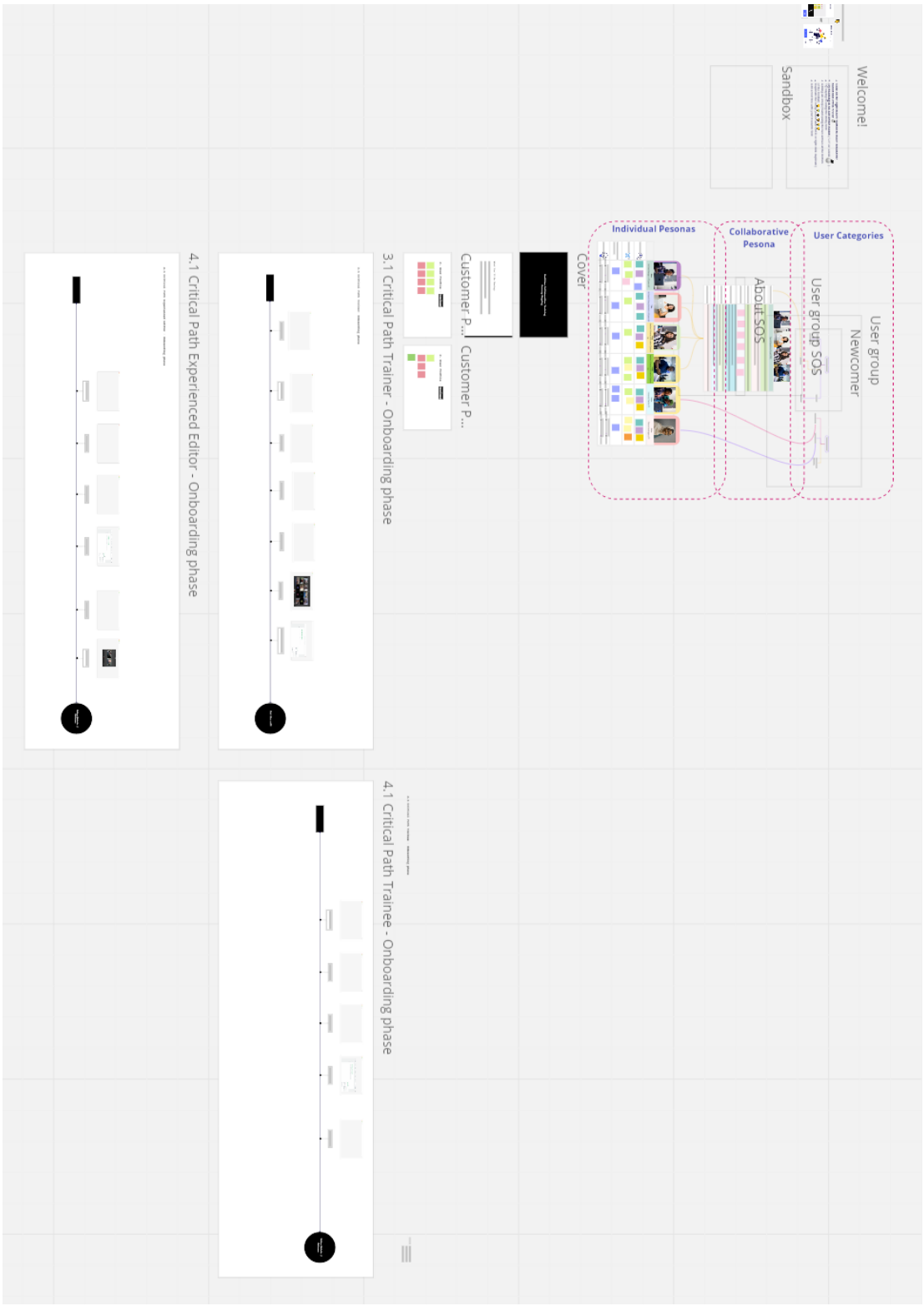
D.5 The Full Miro Board Landscape of the “Position” Session



D.6 The Full Miro Board Landscape of the “Review” Session



D.7 The Full Miro Board Landscape of the “Concept” Session



D.8 Thematic Analysis Codebook

Thematic Analysis Codebook
⮟ Asynchronous collaboration
Asynchronous tasks
Asynchronous editing frustration
⮟ Asynchronous editing challenges
⮟ Avoid having more than one person working on the same article -edit conflict
Workarounds to solve losing text due to edit conflict
Newcomers' assumptions by previous experiences with other collaboration tools
Wiki editing is hard
Word editor of backing up an article that might be deleted
Defining asynchronous collaboration as collaboration
Discussion on getting articles a featured status
⮟ Success factors
Importance of previous coordination for successful async collaboration
Ways that make you feel the collaboration
⮟ Collaboration on Wikipedia
Collaborating with bots
⮟ Collaboration on things other than the writing
Collaborating with trainers to help newcomers
Helping in making the first step
⮟ Collaborative tools
⮟ Internal tools
Organisation pages are useful
Sandbox
Talk page
WikiProject
⮟ The use of external tools to support collaboration
⮟ Familiar

Etherpad for arranging meetings
Facebook
Google
PowerPoint
Twitter
WhatsApp
⮟ Social media to coordinate tasks
Helpful in covid
⮟ Unfamiliar
Etherpad
⮟ Different collaboration model
High volume of edits after an unscheduled event improves the quality
Helping newcomers
Teahouse
Newcomers' challenges
Difficulties in editing Wikipedia using the phone
External newcomer challenge
Finding reliable resources
Internal challenge to the newcomer
Lack of resources
Lack of time for newcomers
Newcomer's lack of confidence
Newcomers' articles might get deleted
Newcomer's lack of clarity
Newcomer's lack of knowledge of how Wikipedia works
⮟ Trainer challenges
⮟ IP block
Solution
It's hard to motivate students to attend training

Newcomers' articles might get deleted
Welcoming newcomers
⮟ Design feature suggestion or requirement
Concerns
⮟ Onboarding experience
⮟ Challenges
Admins rejecting content
Cultural bias
Getting to know the rules
Systematic bias
Understanding the common language and terminology
The first experience of high importance
⮟ Wikipedia training
⮟ Benefits of training
Changing the assumptions about Wikipedia
Developing the trainee skills
Knowledge sharing
Making the learning process easier
Seeing a direct impact on the newcomer's learning
Setting users as confirmed users
Differences between online and physical training
Hybrid training
Importance of easy training
Importance of retaining editors after training
⮟ Need for change
COVID pandemic effect on training models
Raising a hand and asking for help is good for training in physical space
⮟ Trainee group division
Level interest

Level of expertise in Wikipedia
Level of knowledge about the topic
⮟ Trainer goals
Facilitate independence
Trainers' challenges
⮟ Training importance
Newcomers' positive experience with being trained
⮟ Training Motivation
Motivating people to sign up for training through an external tool
Motivation for attending a training
⮟ Training tools
⮟ Pet scan
Getting worklists ready
Trainer preparing articles for translation using tools like pet scan
⮟ Wikipedia training opportunities
Breakout rooms for one-to-one coaching
Tailoring the training according to the trainee groups
Using videos to train people
⮟ Real-time collaboration
⮟ Advertising real-time collaboration events
Informing collaborators
Method to get people to a collaboration session
⮟ Benefits
Benefit of working together in real-time
Having an editing mode that people are familiar with
Sharing information with someone with sources to write about
The collaboration between Wikipedians resulted in changing categories structure
⮟ Challenges

➤ Attribution
Role account problem
Different levels of authorisation
Different levels of skill
Edits getting blocked by admins
Language barriers when collaborating with someone abroad
Life distractions
Too much ownership
Working with someone who you didn't work with before
Current use of real-time collaboration
➤ Demographics
Benefits of having both area enthusiasts and digital expert people
➤ Demographics that would benefit from real-time collaboration
Real-time collaboration for groups who have already done some work previously
Trainee gender in pharmacy course
Training for university students
➤Design considerations
Accommodating the growing percentage of mobile users
➤ Clarity
Having clarity of the main working space and channels objective
Having clarity on who is working on what
Collaboration awareness
Collaboration should involve training people on Wikipedia
Creating a welcoming space
Discussion
Dividing articles into parts, so people collaborate on different sections
➤ Examples and resources to benefit from
Routes and policies
Way to share resources

⮟ Flexibility
Have a backup plan if the collaboration doesn't work
Gradually increasing the difficulty for trainees
Solution to the attribution problem
Suggestions to improve an article by the tool or community
Universal
Usability concerns for real-time editing
⮟ Motivation
Motivation to trainers to use a real-time collaboration tool
⮟Need for real-time collaboration
Need for a new tool or improving existing tools for real-time collaboration
Workarounds to facilitate real-time collaborations
⮟Opportunities
Collaborating on work lists for successful project
⮟Collaborations on lists
Concerns with assigning roles for people on worklists
Connecting with admins to support the process
⮟Content
Less controversial topics for real-time collaboration
Topics to collaborate on
Cooperation
Efficiency
Mobile editing
Motivation to create
Newcomer previous experience with real-time collaboration
Possible methods to collaborate in different languages
Possible ways for collaboration with other editors
Proud memory of stopping vandalism in real-time
⮟ Scenario for collaboration editing

Having a call to support a collaboration
Using real-time collaboration for worklists
Scheduled or unscheduled events collaboration models
Setting expectations for the collaboration
Subject area groups
Support for training from external organisations
Word editor of backing up an article that might be deleted
⮟Potential activities
⮟Worklist
⮟Opportunities
Collaborative alternatives to real-time collaboration
Discussions prior to training
Personal engagement prior to training
⮟Risks
Someone deleting the collaborative document
Two people picking the same article
⮟Real-time collaboration challenges
⮟ Assumptions about a challenge
Assumption that we need to protect an article from editing to do collaborative writing
Assumptions about Wikipedia
Attribution issue of real-time collaborative editing
Coordinating the tasks
Wiki design as an async tool
⮟Risks
Confusing the newcomer with too many communication channels
Disrupted internet connection
Technical difficulties
⮟Roles and distributing tasks among the group

It happens naturally
⮟Problem with assigning roles
Some people might not be able to do pre-assigned tasks
⮟Pros
Achieving the goal collaboratively
Increase confidence
Woking in the area that the trainee is comfortable with
⮟Roles for people in a training
Auditing
Connecting with WikiData
Drafting
Experienced editor-digital helper
Researching
Welcoming environment
Wiki expert moving text to Wikipedia
Trainer role as a facilitator
⮟Transparency concerns
Contradicting the meaning of openness
Importance of openness
Too much transparency causes confusion
Transparency by administrators
⮟Wikipedia challenges
⮟Bias
⮟Cultural bias
Causation of bias
Society's gender inequality in schools
High proportion of unsourced content affects newcomers after reverting
Internet access
Lack of continuation of projects and initiatives

⮟Usability issues
Hard to teach 7th grad how to edit Wikipedia
⮟Wikipedia's community
⮟Challenges
Having computers
Newcomer's lack of knowledge about the rules
⮟Community rules
Changing rules about blocking
Prevent promotional language
Informing the community
⮟Motivation to be involved in Wikipedia
Collective intelligence
Develop good new habits
⮟Develop new skills
Citation
Feeling of belonging
Instant gratification of adding a page
Meeting others from all around the world
Method for forming a neutral, reliable knowledge
The big audience for your writing
Positive experiences

D.9 Participants' Discussion on The IP Block Issue Faced by Nigerian Trainees

After sharing the TA analysis results in the "Concept" session, Participant J asked about the blocking issue, which happens when multiple accounts are created or intensive editing occur in the same location. Participant J explains her concerns as follows:

> "Thank you for this very wonderful presentation. It's very informative, and I think the social challenges are on point, though, to me, I feel like this challenge of people getting blocked. Yes? When they're trying to edit, I think that should

> be one of the social challenges that new editors, even existing editors, face it too. But when trying to edit on Wikipedia or any Wikipedia sister project, they get blocked immediately, and most new editors don't know how to get a solution. So, I'm not sure if this is related, but can you editors get a solution to this blocking issue? It's part of the challenges that we do face also."

This led me to ask other participants to help answer this question, knowing that there are participants with over ten years of experience and different levels of authority in the community attending the session. As a result, Participant A volunteered to answer the following:

> "There's been some discussion on the mailing list recently about one of the problems of blocking and some countries being more effective than others that we have this very long-standing policy of blocking open proxies and VPNs and so on because certain vandals were using that sort of thing to get past other blocks and vandalise from different parts of the world than they really live in. But it's having the effect that certain countries where, they commonly used VPN or have relatively few IP addresses that everything's being channel through are being caught out by this. So, there's a geographic problem. There is discussion going on about it in the community, and we may well have to change the way we're blocking, particularly if we can identify those countries that are being badly affected by this. Now that affects the block, those people who are blocked and they can't edit at all.
>
> There is a different issue where the people have done some edits and been blocked on the English Wikipedia. That tends to be people who are using overly promotional language, so they may be very enthusiastic about this particular band or something, and it comes across as if it's an edit by somebody working for the band rather than a fan or enthusiast who just wants better coverage of that particular footballer or whatever. So, there are two different problems there's an educational one with new editors who are being blocked because they're not writing neutrally, and they're so non-neutral that it comes across as spam."

Participant B entered the conversation and tried to provide a solution by adding:

"Sometimes when I run an Edit-a-Thons, there's a scheme to add individual accounts as a new user (manually confirm new users), and I believe that helped prevent them getting blocked."

Then, Participant A, tried to also find a solution for Participant J by saying the following:

"Yes, there's one of the other filters that we've got is there's a maximum number of new accounts that can be created at one IP address at a time. And there's also a maximum number of edits that can take place from one IP address. Now, these filters are there because sometimes you would have an entire classroom of unsupervised kids going on a vandalism spree. So, the software detects that by saying I've got seven people trying to save an edit in the same few seconds, which is a training session, and you ask the words all to hit save now, you're pretty much guaranteed to have the lesson collapse in a bit of a mess. If you have you have an admin add to your training session, I've done this a few times myself; then having them set new editors there as confirmed editors mean that they don't count towards that filter, and they don't have the really annoying thing of having to do a CAPTCHA every time they're putting a link in [or] citing a source. I think we've got a facility now I forget the name, for effectively whitelisting an IP address that's being used for training professionals, and that's something we discussed a few years ago. I think that now it exists. And again, it solves that problem for training sessions.

There's still the issue of there are certain features of our blocking systems that were kind of sensible 18 or 19 years ago, but don't work if you've got people using mobile phones where the IP address changes all the time. So yes, you could well wind up with the IP address of somebody who was blocked three months ago for making vandalism edits. But if it's one of these IP addresses that keeps moving between every mobile phone. Well, there's no point blocking that IP address for three hours left alone three months [...].

Another get around, particularly with the account creation thing, is that the limit is per wiki, so if you're running a multi-language thing and you've got people editing on four or five different language versions of Wikipedia, they can create

> accounts on each of those. And if you're single language one where you can always get some people creating accounts on a different IP address if you edit on a mobile phone or get them creating accounts on Wikimedia Commons and then editing on Wikipedia."

Participant F, who is also joining from Nigeria and previously pointed out the seriousness of this problem of IP block, added the following:

> "Ordinarily before, I didn't know one could just instead of having all editors creating Wikipedia accounts at the same time, one can just tell them to create accounts on other Wikipedia sisters' projects and use the accounts or and use the accounts that have been created on another Wikipedia sisters' projects to work on Wikipedia [...] here in Nigeria, we're always facing this issue of IP address even without using VPN […] so thank you so much for that idea. I think I would also recommend it to other experience editors when they have issues like this."

Appendix E Social Voting and Ideation

E.1 Online Qualtrics Survey

Prototyping a Collaborative Wikipedia Training and Editing Tool

Participant Information Sheet

We appreciate your interest in our research! We are studying online Wikipedia training and editing in order to prototype and evaluate a collaborative training tool that can support Wikipedia trainers train newcomers to Wikipedia remotely.

The first phase of this research spoke to both new Wikipedia editors and experienced Wikimedians. In this phase, we have a series of suggestions on which we would like the opinions of the wider Wikimedia community.

This survey will let us know how familiar you are with Wikipedia and create a unique ID number for you. At the end of the form, you will be redirected to an online ideation tool for you to view, comment, and vote on others' ideas about designing an online Wikipedia training tool, as well as add your ideas for others to discuss and vote on.

Have you ever edited a Wikipedia article? Please select the answer that is closest to your experience.

- ○ Yes, I have been an active contributor to Wikipedia for ____ years. Please enter the number of years below:
- ○ Yes, I have an account, but I have not been active for more than one year.
- ○ Yes, other
- ○ No, but I have thought about it (please share why you did not edit)
- ○ No, I have never edited or contributed to Wikipedia.
- ○ No, I did not realize that I could.
- ○ No, other

Have you ever attended Wikipedia training?

- ○ Yes, I received a face to face training.

- ◯ Yes, I received an online training.
- ◯ Yes, I received an online training to be a trainer.
- ◯ Yes, I received a face to face training to be a trainer.
- ◯ Yes, Other
- ◯ No
- ◯ No, I have learned how to edit on my own.
- ◯ No, Other

How long has it been since you received the training? (e.g. almost two months, a year)

Who provided you with the training?

- ◯ Wikimedia Chapter
- ◯ Organisation
- ◯ Other

Do you currently have an administrative role in any Wikipedia language version (or any other Wikimedia projects) before?

- ◯ Yes
- ◯ No, but previously I had.
- ◯ No

What is your role? And in which Wikimedia Project/s?

What was your role? And in which Wikimedia Project/s?

Do you currently have a role in Wikimedia Foundation or its chapters?

- ◯ Yes
- ◯ No, but previously I had.
- ◯ No

What is your role? And in which Wikimedia Chapter/Project?

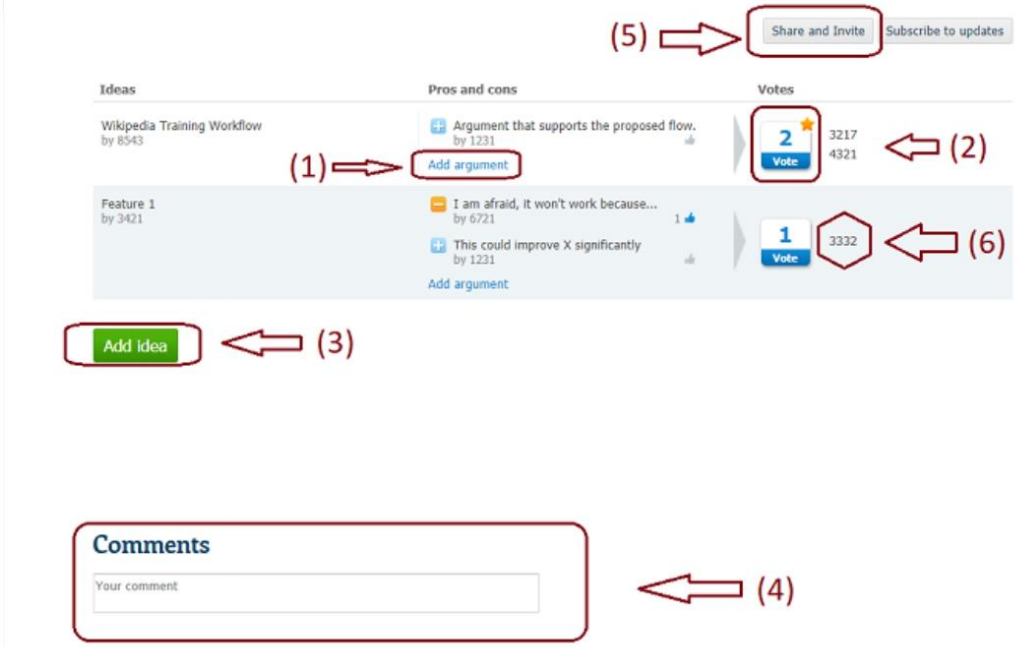
What was your role? And in which Wikimedia Chapter/Project?

Please copy the following number as you will need it to participate in the Tricider social voting tool. **You should use this ID whenever the tool prompts you to add your name, even if the field says that adding a name is optional.**

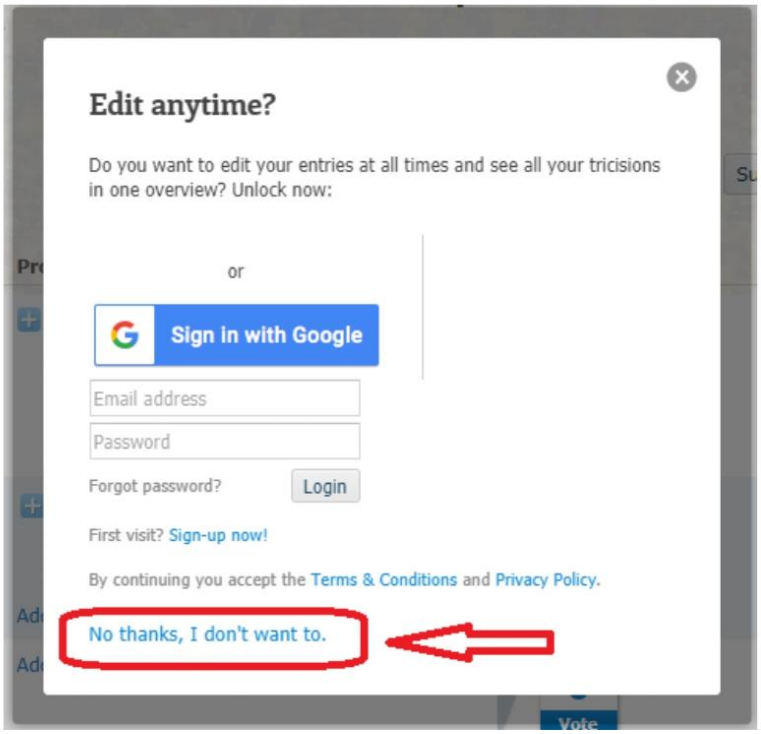
Here is your ID Number:

${e://Field/Random%20ID}

We will be using the Tricider social voting tool.
1- You can add pros/cons arguments to features suggested by participants in earlier research stages.
2- Vote on ideas.
3- Add your ideas, if any!
4- Add your comments, if any.
5- Share this with friends who might find it interesting.
6- Please, connect all of your contributions to your ID number.



You will see this interface the first time you interact with the tool. Please, choose "No thanks, I don't want to" to use the tool without registration.



Ethical Approvals
This research proposal has been scrutinised and subsequently granted ethical approval by the University of St Andrews Teaching and Research Ethics Committee. The approval reference number is CS15836.

What should I do if I have concerns about this study?

In the first instance and before you move to the consent page, you are encouraged to raise your concerns or questions with the researchers and if you do not feel comfortable doing so, then you should contact the research supervisors. A full outline of the procedures governed by the University Teaching and Research Ethics Committee is available at http://www.st-andrews.ac.uk/utrec/guidelinespolicies/complaints
You can revisit the official advertisement from here: Advertisement collaborative Wikipedia training

Contact details

Researcher(s):	Abd Alsattar Ardati aaa8@st-andrews.ac.uk	**Supervisor(s):**	Dr Alexander Voss alex.voss@st-andrews.ac.uk
	Dr Sara Thomas sara.thomas@wikimedia.org.uk		Dr Uta Hinrichs uh3@st-andrews.ac.uk
	Dr Richard Nevell richard.nevell@wikimedia.org.uk		Dr Angela Miguel arm14@st-andrews.ac.uk
	Dr Kirsty Ross ksrh1@st-andrews.ac.uk		Prof Ian Gent ian.gent@st-andrews.ac.uk

Consent Form

I confirm that I am willing to take part in this research.

- ☐ If you accept, please tick the box
- ☐ I confirm that I am over 18 years old of age or older (a requirement by the university ethics)

Please remember to **copy your ID ${e://Field/Random%20ID} somewhere else**, as after you click the next arrow button, you will be redirected to the social voting tool and will not be able to get your ID number.

E.2 Advertisement

The following advertisement channels list was developed and reviewed with the research participants and collaborators, who helped me with dissemination throughout the PD process.

1) Wiki Mailing lists[135]:
    a. Wikimedia UK
    b. Scotwiki mailing list[136]
    c. Wikimedia Research
    d. Wikimedia UK Trainers
    e. EduWiki User Group
2) PSCI-Comm mail list[137]
3) Advertising in newsletters
4) Invitation to the research participants to share the ad among their networks
5) Reaching out to Wikimedia Chapters
6) Wikimedia User Groups
7) Wikimedia Affiliates into the broader Wikimedia community
8) The IDEA Network editors[138]: The IDEA Network[139] is an interdisciplinary group of like-minded staff and students from across the University of St Andrews who work on ensuring that access to knowledge is open to all and reflects the diversity of knowledge in society.
9) ScotPEN Facebook page[140]
10) Industry focused liaison at University of St Andrews.
11) BIG-chat[141]
12) Visitors Studies Group[142]
13) NCCPE email list[143]

[135] https://lists.wikimedia.org/postorius/lists/?count=200
[136] https://docs.google.com/forms/d/e/1FAIpQLSctpnDjO_PHZeOTrA0yjtVLqE4oDfjLtYhWzQVTwk712BSJJQ/viewform
[137] https://scicommjobs.wordpress.com/psci-com-list/
[138] https://outreachdashboard.wmflabs.org/campaigns/open_knowledge__university_of_st_andrews/programs
[139] https://research.wp.st-andrews.ac.uk/2021/04/30/using-wikipedias-force-for-good/
[140] https://www.facebook.com/groups/ScotPEN/
[141] https://www.big.uk.com/bigchat
[142] https://visitors.org.uk/
[143] https://www.publicengagement.ac.uk/connect-with-others/public-engagement-network

From: [redacted]
Sent: 08 November 2022 18:36
To: eduwiki@lists.wikimedia.org
Cc: Abd Alsattar Ardati
Subject: Prototyping a collaborative training tool that can support real-time collaborative writing

Hello everyone,

Abd Alsattar Ardati (University of St Andrews) and collaborators are studying online Wikipedia training and editing to prototype and evaluate a **collaborative training tool that can support Wikipedia trainers to train newcomers to Wikipedia remotely**, and we want to **invite you** to participate in our **research study anonymously**.

If you agree to participate in this study, please click on the following questionnaire link so we can find out how familiar you are with Wikipedia and whether you have any experience in Wikipedia editing. **It will take ~2 mins to complete**.

At the end, **you will be redirected to an online social ideation and voting tool called Tricider** to **view, comment, and vote on others' ideas about designing an online Wikipedia training tool** and to add your ideas for others to discuss and vote on. Please beware of the Tricider website's adverts that would look like part of the listed ideas.

https://standrews.eu.qualtrics.com/jfe/form/SV_8GuEdzPOGFaFrRc

Your contribution is important. Thank you.

If you have any questions, feel free to contact the researchers. Their details are below:

Contact Details

Researchers:
Abd Alsattar Ardati (aaa8@st-andrews.ac.uk)
Dr Sara Thomas (sara.thomas@wikimedia.org.uk)
Dr Richard Nevell (richard.nevell@wikimedia.org.uk)
Dr Kirsty Ross (ksrh1@st-andrews.ac.uk)

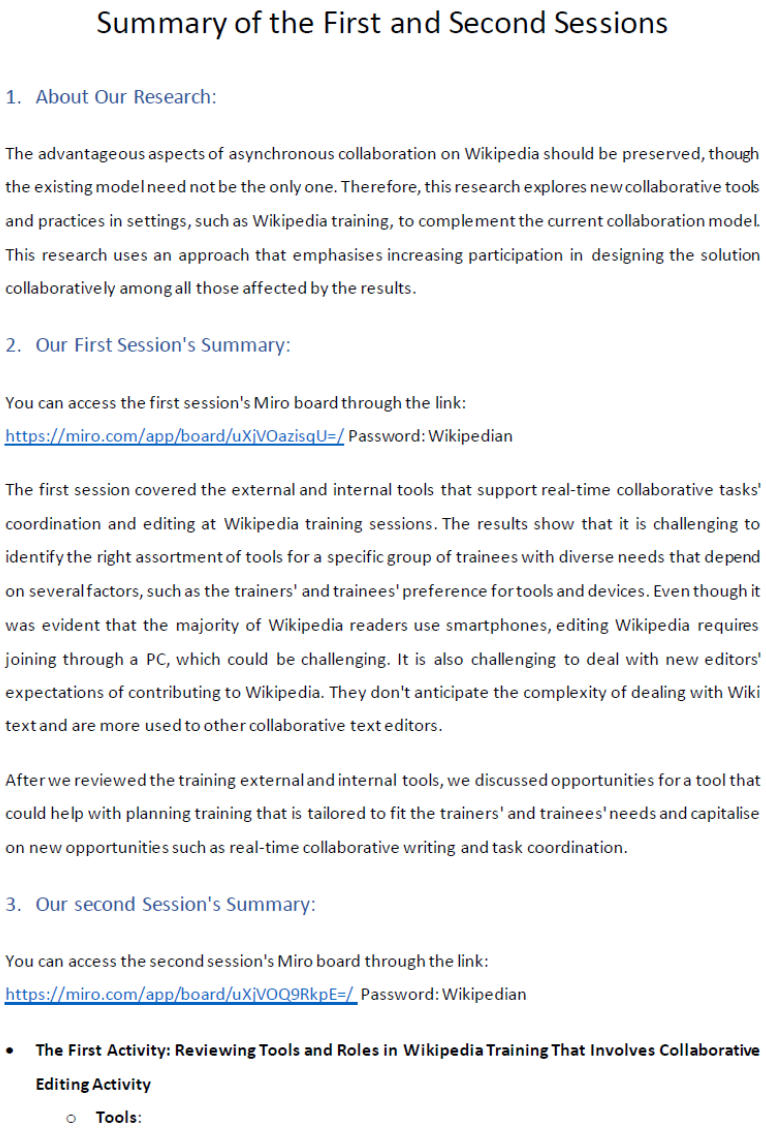
Supervisors:
Prof Ian Gent (ian.gent@st-andrews.ac.uk)
Dr Alexander Voss (alex.voss@st-andrews.ac.uk)
Dr Uta Hinrichs (uh3@st-andrews.ac.uk)
Dr Angela Miguel (arm14@st-andrews.ac.uk)

School of Computer Science
University of St Andrews
Jack Cole Building
North Haugh

Appendix F Reporting to the Participants

F.1 Sessions Summarisation

The first page of the "Identify" and "Define" sessions was sent to participants on the 23rd of Feb, 2022. Most of the content in the rest of the document is covered in the thesis. Therefore, only the first page is included here. Another similar summarisation was sent after the third session.



Summary of the First and Second Sessions

1. About Our Research:

The advantageous aspects of asynchronous collaboration on Wikipedia should be preserved, though the existing model need not be the only one. Therefore, this research explores new collaborative tools and practices in settings, such as Wikipedia training, to complement the current collaboration model. This research uses an approach that emphasises increasing participation in designing the solution collaboratively among all those affected by the results.

2. Our First Session's Summary:

You can access the first session's Miro board through the link: https://miro.com/app/board/uXjVOazisqU=/ Password: Wikipedian

The first session covered the external and internal tools that support real-time collaborative tasks' coordination and editing at Wikipedia training sessions. The results show that it is challenging to identify the right assortment of tools for a specific group of trainees with diverse needs that depend on several factors, such as the trainers' and trainees' preference for tools and devices. Even though it was evident that the majority of Wikipedia readers use smartphones, editing Wikipedia requires joining through a PC, which could be challenging. It is also challenging to deal with new editors' expectations of contributing to Wikipedia. They don't anticipate the complexity of dealing with Wiki text and are more used to other collaborative text editors.

After we reviewed the training external and internal tools, we discussed opportunities for a tool that could help with planning training that is tailored to fit the trainers' and trainees' needs and capitalise on new opportunities such as real-time collaborative writing and task coordination.

3. Our second Session's Summary:

You can access the second session's Miro board through the link: https://miro.com/app/board/uXjVOQ9RkpE=/ Password: Wikipedian

- **The First Activity: Reviewing Tools and Roles in Wikipedia Training That Involves Collaborative Editing Activity**
  - **Tools**:

The following is an example of an email sent to participants with the summary attached to it.

Email example 1:

Abd Alsattar Ardati

From: Abd Alsattar Ardati
Sent: 28 January 2022 16:47
To: Abd Alsattar Ardati
Subject: RE: Second Session: Collaborative Wikipedia Training and Editing
Attachments: Extended_Version_Summary_and_Results_for_the_First_Session.pdf; Short_Version_Results_for_the_First_Session.pdf

Hello again!

As mentioned in the email below, attached please choose which version of our first session's report you would like to read:

a) An extended version for those who want to read the summary and results from the first session. (11 pages, include graphs from the Miro board)
b) A short version for those who wish only to read the results. (3 pages)

Both versions will include an overview of the research and a plan for the next session. This research uses the Participatory Design research approach, which emphasises distributing the decision-making process among all those affected by the solution design. So, I would appreciate any feedback on my interpretation of the first session or the aims of our next one.

Unfortunately, one person found the timing unsuitable but might join us later at the session. I will send Microsoft Teams' invitations shortly after this email.

Many thanks and best wishes,
Abd

From: Abd Alsattar Ardati
Sent: 25 January 2022 11:56
To: Abd Alsattar Ardati <aaa8@st-andrews.ac.uk>
Subject: Second Session: Collaborative Wikipedia Training and Editing

Hello there,

Thank you again for taking part in our first session. It was an enriching experience to have such accomplished trainers and Wikipedians in one place!

Sorry for not following up earlier, I kicked off the new year with some health issues. But, I am doing great now and catching up with research nicely!

In our first session, we focused on Wikipedia's training tools and activities flow. Together we identified the current training model strengths and challenges and discussed opportunities for improving the collaborative editing process.

Our second session will be held **on the 3rd of Feb from 6 pm – 8 pm**. I picked the timing based on your feedback from our first session. If you have a problem with the date or any questions, could you please reply to this email before the end of the week?

You will receive **a Microsoft Teams' invitation and the first session's summary on Friday the 28th**. Please, accept the invitation to confirm your place.

The second session will focus on discussing possibilities, concepts, and goals for a tool that could address some of the Wikipedia training challenges. More details will be found in **the agenda,** which **will be sent one day earlier**.

Many thanks and best wishes,
Abd

University of St Andrews | FOUNDED 1413

Abd Alsattar Ardati
PhD Student
SACHI Group, University of St Andrews, Scotland aaa8@st-andrews.ac.uk | www.st-andrews.ac.uk | abd.ardati.org
phone: +44 (0)1334 46 3260

Email example 2:

Abd Alsattar Ardati

From: Abd Alsattar Ardati
Sent: 03 February 2022 17:02
To: Abd Alsattar Ardati
Subject: RE: Agenda - Second Session: Collaborative Wikipedia Training and Editing

Hello there!

This is a kind reminder for our meeting in an hour at 18:00 pm.

Microsoft Teams: please join the meeting through this link or from the one provided in the invitation email.

During the session, we will use **Miro board** through this link **Password:** Wikipedian

Unfortunately, the meeting clashes with some participants' other engagements. So, some of you may join later or leave earlier.

Please let me know if you have any questions.

Many thanks,
Abd

From: Abd Alsattar Ardati
Sent: 02 February 2022 14:55
To: Abd Alsattar Ardati <aaa8@st-andrews.ac.uk>
Subject: Agenda - Second Session: Collaborative Wikipedia Training and Editing

Hello there,

I appreciate your interest in joining our second session **tomorrow at 6 pm**! Here is our **Agenda:**

1. Welcome.
2. Introducing the study.
3. Round of introductions.
4. Review our first session's findings.
5. Discuss training scenarios that benefit from opportunities covered in the last session to reduce the training challenges.
6. Wrap-up, reflections, and next step.

Please access the session using a laptop or PC to make it easier for you when we move to use the Miro board tool. Also, please, make sure that the tool works on your setup.

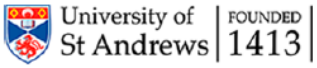
You don't have to create a Miro account; just access it through the following link:
https://miro.com/app/board/uXjVOQ9RkpE=/ **Password:** Wikipedian

Please feel free to try and edit the board as you wish; I will clear it and add our activities before the session. I will introduce the tool again on the session's day.

You should have already received a **Microsoft Team's invitation; please could you accept it, if you haven't done so already**? or let me know if you cannot make it. I will send a reminder with the meeting link one hour before the session.

As always, please let me know if you have any questions, technical difficulties or comments from the last sessions.

Best wishes,
Abd



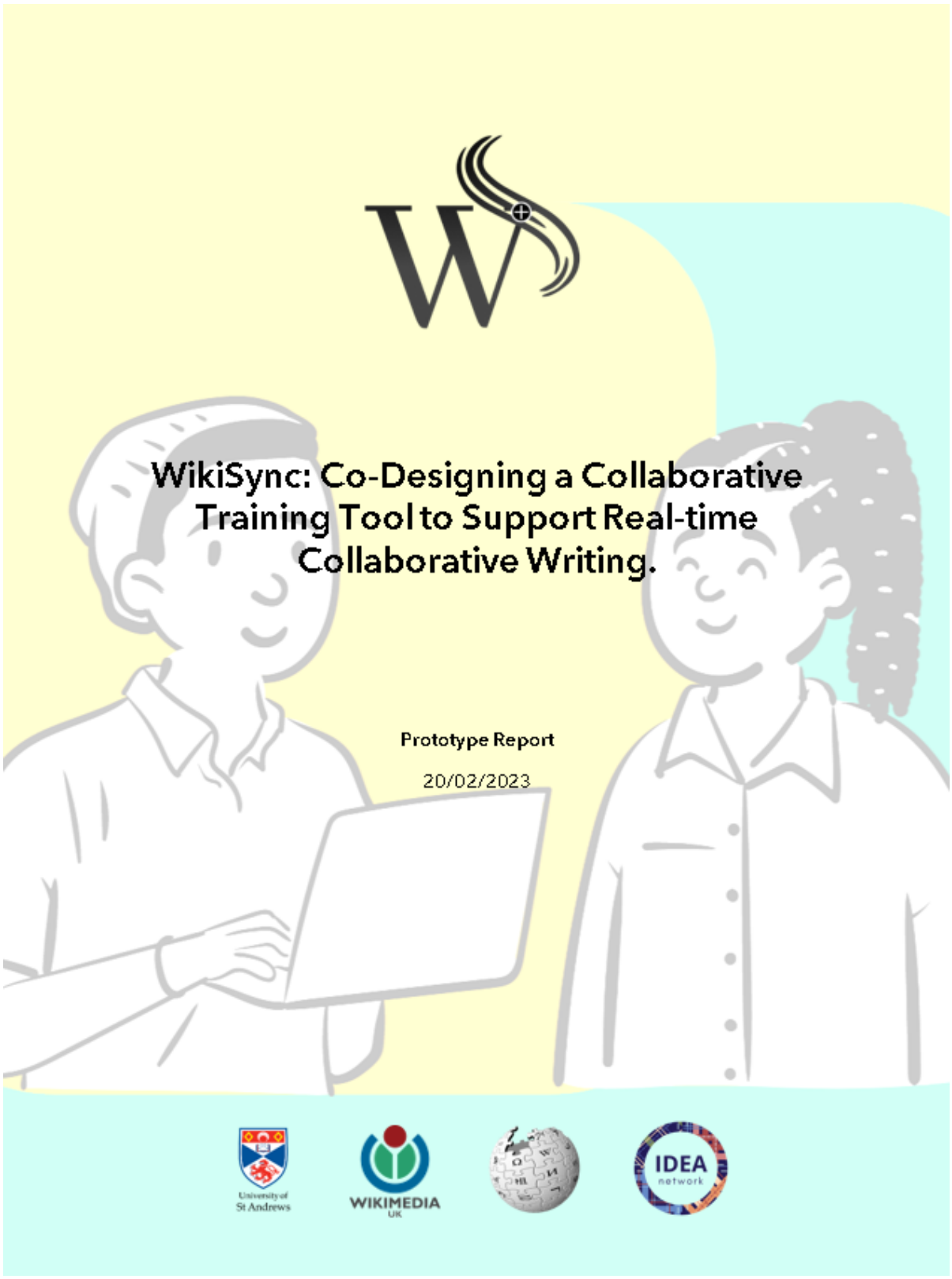
Abd Alsattar Ardati
PhD Student
SACHI Group, University of St Andrews, Scotland aaa8@st-andrews.ac.uk | www.st-andrews.ac.uk | abd.ardati.org
phone: +44 (0)1334 46 3260

I aspire to a healthy life:work balance. Please only respond to my emails during your normal working hours; I do not expect a response outwith these hours.

F.2 Prototype Report

The following two figures show the first two pages of a 35 pages report sent to the participants before the last session. The full report can be found through the following link.

https://standrews.eu.qualtrics.com/WRQualtricsControlPanel/File.php?F=F_TPmKqKY79RSwFHN


WikiSync: Co-Designing a Collaborative Training Tool to Support Real-time Collaborative Writing.
Prototype Report
20/02/2023
University of St Andrews
WIKIMEDIA UK
IDEA network

Contents

Appendix G EDPD Framework Guide

Welcome to the framework guide, which is written to walk you through an Ethnographically-informed Distributed Participatory Design (EDPD) Framework to respond to the needs of small teams dealing with large communities. The steps in this guide provide a short, intelligent description on using the framework, highlighting the key points to aid users in applying and adapting the framework to their use cases.

G.1 This Guide Will Help You:

1) **Improve your online design practice in complex social environments.** A complex social environment that facilitates mass collaboration requires new innovative approaches that involve the community in finding software solutions. The project's complexity means that tracking issues and accounting for design decisions are particularly significant.
2) **Achieve high impact.** Rooting the process within the community ensures identifying the problems and their consequences which designing solutions for can have the highest positive impact on the community.
3) **Gain high stakeholders' acceptance.** The framework's phasic approach to ensuring a diverse community representation in co-constructing the IT platform provides higher levels of solution acceptance.

G.2 Who Is This For?

This novel design approach has the needs of small teams in mind. Are you dealing with a complex change case in your organisation? Is the community in focus scattered geographically? How, who, and how many to engage, and at which stage? Multiple questions emerge when working with products in our ever-evolving, increasingly complex world.

This guide is intended for individuals (facilitators) researching a specific change or design solutions for a focus area within the community. It helps guide those leading the implementation of such an effort through a novel framework for active community participation.

Below are four contexts. Applying the framework in the first context has excellent potential in producing solutions with high acceptance and success, while the second, third, and fourth contexts require further adaptation to fit the framework and paying attention to the politics in introducing the change.

1) Non-profit platforms run in partnership with a community of volunteers. Applying this framework can be valuable to the OpenStreetMap (OSM) community,[1] Translators without Borders[2] and the Khan Academy Support Community.[3] Furthermore, community-run software projects that aim to establish platforms, such as Mastodon[4] or PeerTube[5] face similar issues, as they must adapt to increased uptake by people whose expectations are shaped by corporate platforms.
2) For-profit platforms run in partnership with a community of volunteers: This applies to community-led for-profit platforms, such as Reddit and Stack Overflow. However, it is contingent on having an empowered community and corporations that are supportive of their communities.
3) Brand communities: This framework supports facilitators considering a new user-inclusive innovation approach in the business and for-profit environment. While it can be used to explore solutions with these corporate communities in the brand periphery, such as the LEGO Ideas community[6] or Apple Support Community,[7] using it requires careful consideration of the power dynamics involved.
4) People-centric organisations: This framework can benefit organisations seeking a decentralised decision-making environment and building an employee-community-driven culture. It aids User Researchers and Product Managers' understanding of context better, ahead of bringing cross-functional teams together with users in developing their products. Similar to the second and third contexts, this one also demands paying great attention to power dynamics and process transparency, as these are key elements in building the trust needed to get participants' ownership, as successful solutions are devised.

In all organisational contexts, the leader in facilitating the implementation of the framework could be a Community Relations Specialist, User Researcher, Product Designer, User Experience (UX) Designer, or Interaction Designer.

[1] https://community.openstreetmap.org/
[2] https://translatorswithoutborders.org/
[3] https://support.khanacademy.org/hc/en-us/community/topics
[4] https://joinmastodon.org/
[5] https://joinpeertube.org/
[6] https://ideas.lego.com/community
[7] https://discussions.apple.com/welcome

A Product Manager (or an equivalent leading role) should be heavily involved in supporting the facilitator's planning and implementation, especially when the project focus is entangled with the organisation's politics. It is important that the Product Manager and any identified experts in the focus area be available to provide expert reviews of the outputs, especially when moving from one stage to another.

G.3 Process Overview and Facilitation Tips

As shown in Figure G.1, the top and bottom of the process, in parallel with conducting the project, two key activities supplement the process:

- Reaching targeted community groups using relevant communication channels, such as a mailing list or a Facebook group, tools, such as Microsoft Office or a WhatsApp call, and methods, such as online group meetings or social voting.
- Researching and designing suitable collaboration and analysis activities, which this guide unravels throughout its phases.

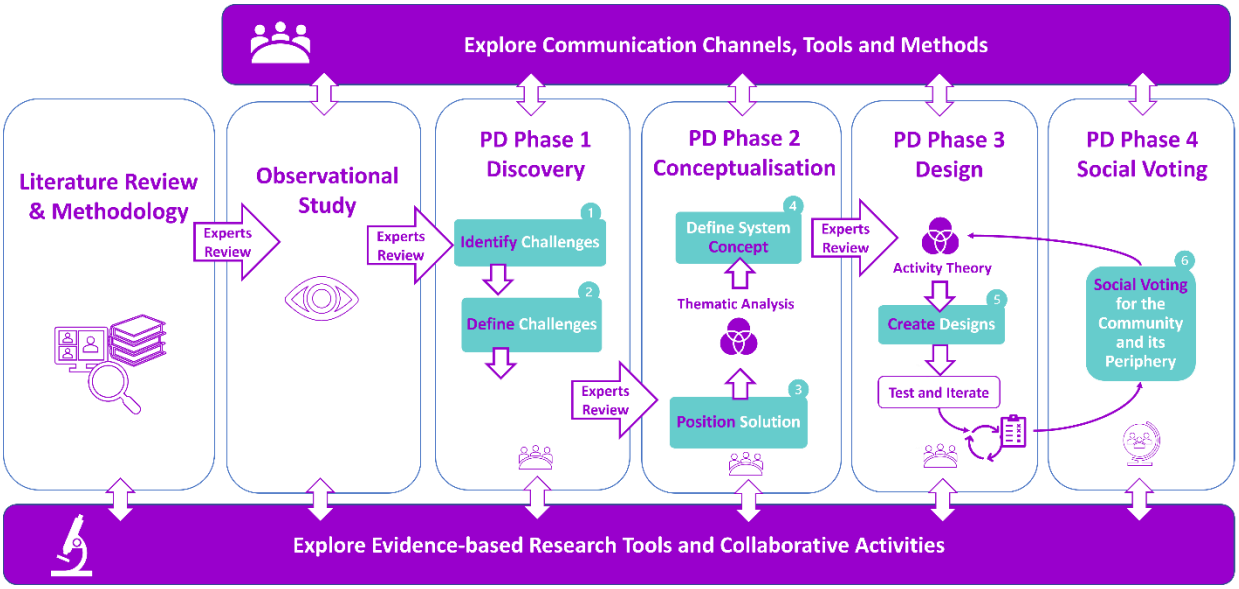


Figure G.1 The EDPD framework phases and sessions.

The EDPD framework is divided into several phases.

1) **Literature Review & Methodology,** which has two parts:

A) Initiating

This phase defines the initial scope and questions that lead to the initiation of the process, which could be triggered by a community's need for change or their demands in the organisation's

innovation process. This provides insights leading to questions about the challenges and potential of introducing changes to such areas as services, products, moderation, structure and engagement. The source data for initiating the process could range from quantitative insights based on data analysis to the strategic direction to qualitative data, such as user feedback, to a combination of both in the form of reports and recommendations from the User Researcher or someone with deep knowledge of the industry and its users.

At this early stage, it is important for the facilitator to be reflective and open about the possibilities, as well as try to balance designing the initial questions based on assumptions about the reasons leading to the challenge. Expert views and reviews are important at this stage.

B) Developing the initial focus

Defining the initial scope of the question about the challenge or requested change. This should help define the areas of literature and industry practice, if any, that could help form an initial understanding of the underlying factors that led to this question. This should help define where and when to observe the community in order to form a better understanding of addressing the question and outline how to approach and arrange ways to observe the community and participate in their activities. An initial list of community tools based on the knowledge gained from this stage could help define how to engage with the community through methods they are familiar and comfortable with.

The initial focus need not be meticulous, rather, it should help one find a place to start with community participation. Keeping a lean mentality towards the scope is essential, as findings emerge from the observation and early stage of the PD process. However, being too agile could lead to scope creep. Defining the scope and estimating its length rely on the change in focus and the number of people involved. Therefore, it is important to utilise project management techniques in defining the time, scope and cost associated with the project at this early stage.

2) Observational Study

The study of the community within their environment and understanding how the area of focus activities are accomplished is essential for designing the sessions that directly engage the end users in the design process through information gathered from their practices. This phase should involve:

- Listing the community tools for ecosystem mapping to understand their role in the community norms and traditions in the area in focus.

- Defining community roles.
- Initial journey mapping of the area in focus.

Even though this stage may result in early design insights, the observation should focus on documenting what is there, rather than discovering what is missing.

3) PD Phase 1 Discovery

This phase involves bringing users along with designers to discover the challenges from the community perspective and understand the context of impactful solutions. It includes two co-design sessions: “identify” and “define”, used as a discovery tool in which methods and techniques can help map the challenges and opportunities and understand the context. For example, these sessions may involve collaborating with the community on exploring the focus area through:

- Mapping the tools used.
- Listing their activities.
- Defining the user groups for the area in focus.

In this phase, participants should be encouraged to collaborate with the facilitator on grouping and deciding what is relevant for further analyses, helping her create outcome summaries that gain everyone’s acceptance before moving on to the next session or phase. This helped participants build ownership over the process and confidence in their capabilities for not only producing data but also analysing it.

- This involves activities that possess critical thinking and analysis, such as:
- Collaborative activities clustering and mapping (Affinity Diagramming).
- Reverse brainstorming about the change.
- SWOT analysis.

It is important to gather as much information as possible about the community’s challenges and aspirations, instead of focusing only on the required change. It is equally crucial to ensure the diversity of the participants, which may result in involving new participants as the process moves forwards.

4) PD Phase 2 Conceptualisation

In this phase, participants collaborate on consolidating a system solution concept, bringing focus to defining the research direction and expected outcomes through engaging diverse representation of the community. This phase involves two co-design sessions: “position” and “concept”, used as

a conceptualisation tool to engage with the community in a process that leads to a short statement outlining the core system concept and the project's scope.

Collaborating with the community on a system statement as an outcome of this phase involves the following activities producing a short and clear definition of the possible solution:

- Defining collaborative persona.
- Collaborative journey mapping.
- Scenario co-designing.

This phase may require onboarding new participants to validate the research direction and diversify the voices. It also may witness an accumulation in the data to understand and identify the emerging themes while exploring the challenges discussed in the Discovery phase.

If the generated information is too voluminous to analyse and extract insights, using tools such as thematic analysis can help translate it into something that would drive the process forward. After defining the system statement, the level of community participation could be reduced to only providing feedback. This could be helpful, especially for small teams with resource constraints. When the technical expertise needed to produce certain artefacts is high, some of these activities are best executed by the facilitator and then discussed and iterated with the community in the PD sessions.

5) PD Phase 3 Design

This phase is about transforming the concepts from the Conceptualisation phase into prototypes and materialising the design of the system through an iterative process that involves the sessions participants and the global community. This phase involves iterative "create" sessions, including prototyping activities through online virtual board tools or sketching with a pen-and-paper method.

The following activities can foster collaboration with the community on designing the solution and iterating on the design up to the point where participants are satisfied and can verify the validity of the resulting solution for their community:

- Storyboarding.
- Business process mapping.
- Sketching.
- Prototyping.

If the thematic analysis from Phase 2 uncovers many socio-technical challenges, analytical lenses such as Activity Theory can help derive insights and design requirements for the new system that would energise the process.

Although most of these activities can be done on virtual boards, such as Miro, synchronously, keeping the asynchronous feedback option is important. For example, someone could email a pen and paper sketch of an interface after the session.

6) PD Phase 4 Social Voting

This phase involves inviting the wider community to comment, vote on other tool design ideas, and add new ideas for all to discuss and vote on using an online social voting tool. However, before involving the broader community at an early stage in long debates on whether a change could be implemented, one must address the most critical concerns of the community members involved in the PD sessions.

Once the solution has matured, the broader community should be consulted through ideation, voting, and debating tools using the appropriate communication channels for an additional validation stage. It also invites the wider community to comment, vote on other tool design ideas, and add new ideas for all to discuss and vote on using an online social voting tool. If the thematic analysis from Phase 2 reveals many socio-technical challenges, analytical lenses such as the activity theory can help obtain insights and design requirements for the new system.

Once the broader community feedback is analysed, and if the community results show significant concerns or suggestions for improvements that challenge the emerged designs, it would be imperative to collaborate with the participants to address those suggestions.

7) PD Phase 5 Use

Finally, building the tool and the system to be tested by the community is an important stage to validate the solution in practice. The approach of this stage should keep the participatory method and may involve revisiting any earlier stage for consolidation. This stage can include activities and methods, such as:

- A/B testing.
- Quantitative surveys.
- A second stage of affinity diagramming with the community.
- Qualitative and quantitative data analysis.

Figure G.2 illustrates the proposed EDPD framework components while combining the methods and techniques used throughout this thesis activities. Activities are outlined approximately in parallel to the phase in which it was used, yet loosely to preserve flexibility and invite other researchers or professionals to adapt it to their needs. The letter (G) next to the activity signifies that it is recommended to divide the working participants into subgroups. The letter (A), on the other hand, indicates the activity's suitability for an asynchronous form of participation. As shown, iteration of the phase should take place only in case of unaddressed concerns that increase in velocity in the last three.

The smaller the team, the more helpful it would be to have an expert or advisory panel with the necessary expertise to consult on the process, starting from the early stage of defining the initial questions to implementation. Their participation would vary based on such factors as expertise, availability, and how critical the situation at hand is. Such consultation could be valuable in designing the workshops and activities or reviewing the results of the analysis stage. However, involving them in all details could slow the project and may result in a loss of momentum. Therefore, it is highly recommended that they have a balanced involvement with the minimum of helping the facilitator with the decision to move from one phase to another.

Once participants finish the session, it is important to keep an asynchronous channel open, such as email communication or a space dedicated for feedback on the session's Miro board, to get their feedback on the session activities and any produced results, in addition to the decision to move from one phase to another.

The diagram shows an additional PD Phase 5 Use that covers the use and evaluation of an implemented system. While this is not covered in this thesis, it is a logical extension of the phasic model. Activities and methods that can be utilised in this phase are shown similarly to other phases. The dotted arrows in PD Phase 5 Use denote that receiving feedback and experimenting with the tool may result in revisiting Phase 3 Design and iterating with the participants on new versions to be tested in Phase 5. If the changes are significant, this may need yet another consultation with the broader community in PD Phase 4 before moving to Phase 5 again.

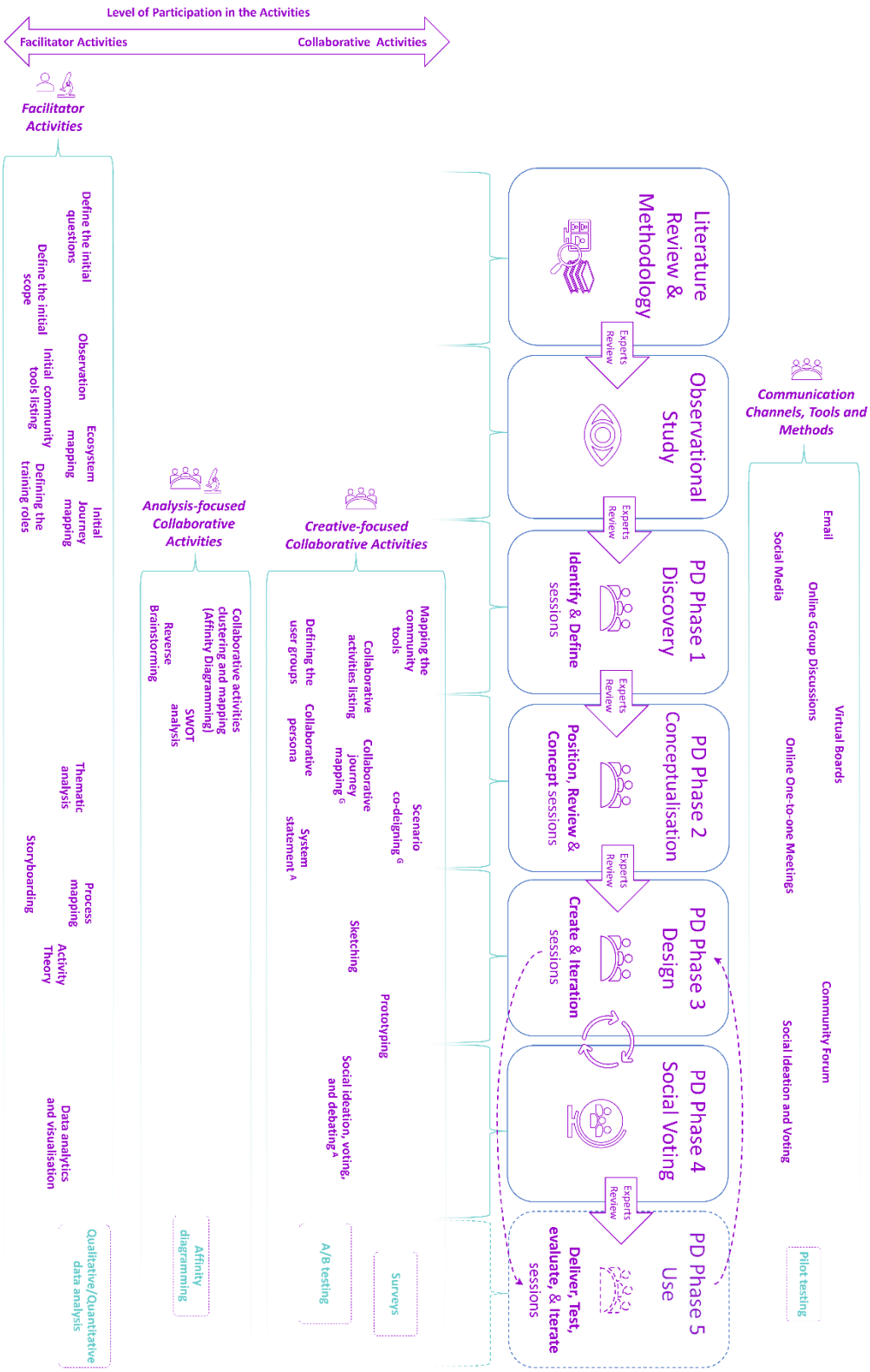


Figure G.2 Ethnographically-informed Distributed Participatory Design framework.

G.4 EDPD Underlying Principles

This framework adopts the principles of participatory design and ethnography that emphasise understanding the context of users and ensuring direct engagement in shaping the systems. Before starting the PD process, facilitators must have a solid understanding of community norms, culture and historical evolution through participating in community activities, consulting experts, or using other necessary methods.

In addition, below are six principles that should tune the relationship between the facilitator and the participants:

1) Face-to-face co-design should be prioritised; however, if that is impossible or would lead to exclusion, conducting online PD should be recognised as an opportunity, assuring community representation.
2) Consulting with field expertise whenever possible. There is never too much consultation with expertise, especially before starting the PD process and moving from one phase to another.
3) In the PD process, early community involvement should not be limited to specific projects that the facilitator thinks need such involvement. Rather, changing a feature that affects the community norms should be thought up starting from the Discovery phase in a small inclusive group, leading to the stage where the broader community is consulted.
4) In designing solutions that serve a large, widely distributed community, it is important to evaluate the right time scale up and scale down the participation, considering factors such as the solution maturity and readiness for broad debates and time and technology available to the participants' diverse backgrounds. Therefore, it is vital to assess and find the proper phase and size of the crowd in which their participation in the innovation process could drive it forward rather than cease its progress or direct it for the benefit of one group over the other.
5) Taking into account the community ecosystem. Efforts to meet long-term community challenges should examine the interplay of the different socio-technical components forming the community ecosystem, starting with forming an understanding of how its technical infrastructure nurtures its community norms and how contemporary technological solutions can complement its ecosystem.

6) The PD facilitator is tasked with being a trust builder and motivator as well as designing the right question based on the accumulated knowledge gained from the session, directing the conversation in the right direction. Additionally, the facilitator should bring value through her broad exposure to literature and expertise in the field and advocate to ensure no one is left behind. Most importantly, the challenging task of knowing when to step back and become an observer, leaving the conversation to the participants, knowing that limited intervention would disrupt a conversation flow that might result in an interesting insight.