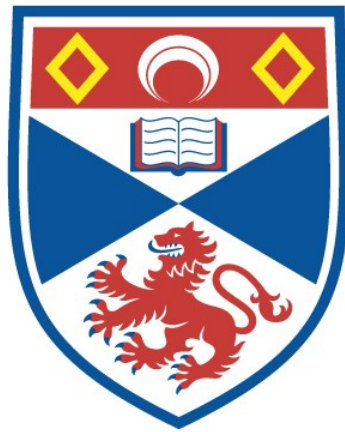# Exoplanets in science fiction

Emma Puranen

A thesis submitted for the degree of PhD
at the
University of St Andrews

2024

Declaration

Candidate's declaration

I, Emma Puranen, do hereby certify that this thesis, submitted for the degree of PhD, which is approximately 39,000 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 September 2020.

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

Date   Signature of candidate

06/08/2024

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   Signature of supervisor

06/08/2024

Underpinning Research Data or Digital Outputs

Candidate's declaration

I, Emma Puranen, understand that by declaring that I have original research data or digital outputs, I should make every effort in meeting the University's and research funders' requirements on the deposit and sharing of research data or research digital outputs.

Date   Signature of candidate

06/08/2024

Permission for publication of underpinning research data or digital outputs

We understand that for any original research data or digital outputs which are deposited, we are giving permission for them to be made available for use in accordance with the requirements of the University and research funders, for the time being in force.

We also understand that the title and the description will be published, and that the underpinning research data or digital outputs will be electronically accessible for use in accordance with the license specified at the point of deposit, unless exempt by award of an embargo as requested below.

The following is an agreed request by candidate and supervisor regarding the publication of underpinning research data or digital outputs: No embargo on underpinning research data or digital outputs.

Date   Signature of candidate

06/08/2024

Date   Signature of supervisor

06/08/2024

Abstract

Science fiction (SF) is a genre influenced by science which, in turn, influences science. Astronomers discovered the first exoplanets (planets outside our solar system) in the 1990s, but exoplanets featured in SF long before then. I aim to answer the research questions: what is the relationship between exoplanet science and science fiction worldbuilding? Can SF be used for science education? This thesis employs an interdisciplinary mix of quantitative and qualitative methodologies to study the relationship between science and SF with a focus on the portrayal of exoplanets. First, I compile a database of 212 SF exoplanets and use Bayesian network analysis to find how various characteristics of these fictional exoplanets influence each other. Results indicate fictional exoplanets created after the discovery of real exoplanets are less Earth-like, evidencing that SF changes in response to scientific discoveries of inhospitable exoplanets. Second, I collect and analyse questionnaire data from participants in my project to create short SF stories in teams of one scientist and one writer. Results show scientific concerns were incorporated into the story creation decision-making process, and suggest an inspirational role for SF towards its readership's interest in science. Third, I interview exoplanet scientists about the relevance of the N=1 Problem, the fact that we know of only one planet with life, on their work. Analysis of interview results shows a diversity of methods to manage the N=1 Problem in exoplanet science research, as well as a trend within this set of participants of N=1 being not formally taught but rather discussed in less formal academic settings. Through the use of quantitative and qualitative methodologies, I provide statistical evidence that SF incorporates rapidly-evolving exoplanet science discoveries into its storytelling. I identify a role for SF in science education of introducing concepts and inspiring general audiences to pursue further study of science.

Acknowledgements

General acknowledgements

The old "it takes a village" adage is true for theses, too. There are many people I would like to thank for their support over the last four years.

I would like to thank my supervisors, Emily Finer, Christiane Helling, V Anne Smith (and Ian Bonnell, for stepping in when Christiane took her exciting new post in Austria). An interdisciplinary thesis presents unique rewards and challenges, and you have all seen me through thick and thin and provided excellent feedback and advice along the way. I am forever grateful for your adaptability and faith in this project.

To my office-mates past and present in the Astronomy PhD Office, thank you for every Cake Tuesday, "Friday" "Lunch", and questionable piece of art drawn on the whiteboard. Having a wonderfully supportive office environment was a huge help along the way.

A few other shout-outs, in no particular order: thank you to Kate, Alex, Kai, and Anna for following me out to Scotland for a visit. Thank you to the void—you know who you are. Thank you to the entire crew of the *RRS Discovery*, past and present. Thank you to Clara, for everything.

Finally, I am always grateful for the support of my family, who inspire and support me to do anything I set my mind to!

Research Data/Digital Outputs access statement

Research data underpinning this thesis are available at:

- https://doi.org/10.17630/64ce9540-7c2b-4794-ace3-5d10b8f0d91c (data set of fictional exoplanets)

- https://doi.org/10.17630/e77b3d39-282c-4900-b89b-273a4c28654a (data set of anthology questionnaires)

- https://doi.org/10.17630/5982566e-8d7b-4a69-87ac-c40c31189539 (data set of N=1 Problem interview transcripts)

Contents

List of Figures

List of Tables

Introduction

This thesis studies the portrayal of exoplanet science in science fiction (SF), using quantitative and qualitative methodologies, with the aim of investigating the relationship between science and SF and the potential uses of SF in science education. The research questions central to this thesis are the following: *What is the relationship between science and science fiction?* and *Can science fiction be used for science education?*. These questions are addressed by focusing on the portrayal of exoplanet science in science fiction, because exoplanets have significant representation in the genre and the field of exoplanet science has seen recent and significant discoveries that have changed human knowledge of planets outside our solar system. Through three research projects, I show that SF incorporates rapidly-evolving scientific discoveries into its storytelling, and identify a role for SF in science education of introducing concepts to and inspiring audiences to pursue further study of science. This role can be expanded incorporation of SF into curricula. I also introduce Bayesian networks as a methodology to supplement close reading to study interrelationships in media datasets. This is an interdisciplinary thesis, and as such blends methodologies of science and humanities theses both.

Chapter 1 provides a background on exoplanet science, the subfield of astronomy engaged in finding and characterising planets outside our solar system. Exoplanets were first confirmed observationally in the 1990s, and in the subsequent decades have become a major area of research (Mayor and Queloz (1995)). Detection methods used to find exoplanets are often indirect, focusing on observing the star, due to the relative dimness of the exoplanets

themselves. Over 5,000 exoplanets have been discovered, and many of them do not fit into recognisable categories from our own solar system: for example, some are gas giants orbiting very close to their stars, others are intermediary in mass between Earth and Neptune (NASA Exoplanet Archive (2024)). While finding an Earth-twin (an Earth-mass planet orbiting at one Earth-Sun distance from a Sun-like star) is beyond current technical capabilities, ongoing missions like the James Webb Space Telescope allow astronomers to identify chemical species in the atmospheres of some exoplanets (Beichman and Green (2018)). Upcoming missions like PLAnetary Transits and Oscillations of stars (PLATO) and Atmospheric Remote-sensing Infrared Exoplanet Large-survey (ARIEL) will further increase ability to characterise exoplanets, increasing our chemical understanding of their atmospheres and searching for potentially habitable worlds around Sun-like stars (Rauer et al. (2014); Tinetti et al. (2018)). Exoplanet science missions provide pieces of data about exoplanets, and theoretical astronomers fill in the gaps using modelling. For centuries before exoplanets were observationally confirmed, science fiction (SF) authors have been writing stories about them. SF is a genre that has basis in scientific fact, and SF stories about exoplanets are also attempts to fill these gaps in the known science of exoplanet data. I conclude the chapter by asserting that exoplanet science is a rich topic with which to explore the portrayal of science in science fiction. This is because exoplanets have long featured in SF and our scientific knowledge of them has recently greatly increased; therefore exoplanets offer a means with which to investigate whether SF responds to scientific discovery.

Chapter 2 defines science fiction as a speculative genre with a basis in scientific fact, featuring cognitive estrangement of the reader's empirical world (Suvin (1979)). I characterise the complex relationship between science fiction and science fact as a reciprocal relationship of mutual influence, with support from historical analysis, author interviews, and case studies. Next I introduce the idea of science fiction's use in science education, addressing current uses like university courses while providing an overview of scholarship in this area which is largely focused on SF audience reactions and understanding of science (Luokkala (2014); Saunders et al. (2004); Kane (2023)). I argue that, due to evidence for the usefulness of narratives in education, SF stories are worthy to consider for science education; however, I conclude by recognising a need to analyse the scientific content of SF to determine its use for science education. I argue that exoplanet science provides an ideal dataset for such a study because exoplanets have long been presented in SF and have recently seen large increases in our scientific

understanding, especially since the 1995 discovery by Mayor and Queloz of the first exoplanet orbiting a Sun-like star. In the tradition of the scientific study of literature by scholars like the Formalists, this thesis analyses a larger dataset of science fiction works to understand their portrayal of exoplanet science.

Chapter 3 provides the methodologies, which I developed, behind the research results presented in Chapters 4, 5, 6, and 7. Firstly, I collected a database, analysed in Chapter 4, of fictional exoplanets from science fiction works, measuring nine variables capturing characteristics for each fictional exoplanet such as whether they orbit a real star or are home to life. I analysed this database using a Bayesian network, a probabilistic graphical network composed of the variables and links between variables that statistically influence each other (Smith (2010)). This network can be analysed to determine how exoplanets are presented in science fiction, which is done in Chapter 5. Another project collected qualitative interview data relating to the decision-making process in creating SF stories in the creation of an anthology through the St Andrews Centre for Exoplanet Science (StACES). The goal of this project was to determine the relationship between science and fiction in the writing process when a science consultant is involved. These results are presented in Chapter 6. Finally, the N=1 project collected qualitative interview data from exoplanet scientists about the effect of the N=1 Problem, the fact that we only know of one planet with life, on exoplanet science. This study was motivated not only by the consideration of the N=1 Problem in exoplanet science but also by the effect of having one example of life on SF worldbuilding. Results are analysed in Chapter 7. This thesis employs an interdisciplinary mix of quantitative and qualitative methodologies to study the relationship between science and SF with a focus on the portrayal of exoplanets.

Chapter 4 presents my database of 212 exoplanets in science fiction. Several examples are given of how planets were categorised according to my nine variables, which define characteristics related to the planets' compositions, stars, orbital locations, habitability, and atmosphere type, as well as whether the planets appeared in fiction before or after the 1995 discovery of real exoplanets, and what type of fictional medium they appeared in. Overall statistics of the fictional exoplanets in the database are compiled, which reveal that 59% are terrestrial planets in the habitable zone (defined as the distance range from a star at which a planet can have liquid water on its surface given an Earth-like atmosphere; see Chapters 2 and 4 for more) with human-breathable air. A comparison with mass and orbital period data from real exoplanets demonstrates that to date, no Earth-twin exoplanet has yet been discovered

(in terms of mass, orbital period, and classification of the host star), and the discovered population of exoplanets is much less Earth-like than my fictional database. Potential reasons for this include narrative demand: SF stories often feature human characters who choose to visit comfortable planets, and budget demand: visual mediums are more restricted in the types of planets they can portray, as well as the influence of the N=1 Problem (see Chapter 7) on science fiction worldbuilding. SF planets often invite a comparison by the audience to Earth, and the differences constitute a form of planetary characterisation—planet-as-character—that allows the environment to play a major role in the storytelling. The chapter concludes with a consideration of how fictional exoplanets can reflect the contemporary science of the time in which they were written, and how the database represents a repository of human understandings and imaginings of other worlds. The database itself therefore represents a result, in addition to the Bayesian network result presented in the next chapter.

The Bayesian network of the fictional exoplanets database is shown in Chapter 5. The network includes eleven links between variables. There is a strong cluster of Earth-like characteristics (human-breathable atmosphere, life, habitable zone, etc.) that positively influence each other, demonstrating both the prevalence of Earth-like fictional exoplanets in the database and that authors combine sets of Earth-like characteristics in worldbuilding their fictional exoplanets. The network also indicates that fictional exoplanets from after the 1995 discovery of real exoplanets are less likely to be home to intelligent native life, and less likely to be home to established populations of non-native human settlers. These results are consistent with findings from exoplanet science of exoplanets that are very inhospitable to human life, or indeed to life as we know it. The Bayesian network thus provides quantitative evidence that the portrayal of exoplanets in science fiction does respond to scientific discoveries. Potential uses of the network to guide science education efforts, as well as informed proposals for explanations for the links such as increasing awareness of the interconnectedness of ecological systems, are made. Bayesian network methodology provides quantitative rigour to the endeavour to understand the portrayal of science in SF, and results show that SF is responsive to changes in scientific knowledge.

Chapter 6 presents and analyses the results of the St Andrews Centre for Exoplanet Science anthology project. I initiated the creation of a science fiction anthology (collection of short stories), entitled *Around Distant Suns: Nine Stories Inspired by the St Andrews Centre for Exoplanet Science*, as well as a research component to understand the role of science in the decision-

making process in SF worldbuilding. Each story in the anthology was written by a team of one scientist and one creative writer, with the story having a basis in the scientific research of the scientist team member. The scientist was to play a significant role in story generation, modelled on the contributions of science consultants to SF stories, but more involved than the typical science consultant. Participants filled in questionnaires asking about story ideas, decisions, and communication between team members. The qualitative data were analysed and categorised by common themes. Results suggest participants saw the role of science in the stories as providing a jumping-off point, a source of inspiration to readers to further explore scientific questions themselves. Participants also embraced the unknown in their stories, reflecting on the unknown in science. Results must not be generalised beyond the small sample size and unique circumstances of this interdisciplinary collaboration, but indicate a potential inspirational, exploratory role of SF in science education efforts.

Chapter 7 presents the results of a project investigating the effects of the N=1 Problem, the fact that we only know of one example of a planet with life, on exoplanet science. 18 exoplanet science researchers were interviewed about their awareness of the N=1 Problem and its impact on their research. Results found that while all were aware of the concept, none were formally taught it and few had heard it referred to as the "N=1 Problem", indicating that the N=1 Problem might be prevalent but not often discussed in academic settings in exoplanet science. Participants were varied in their responses about the impact of the N=1 Problem on exoplanet science, with some finding it limited research to narrow possibilities for life and others responding it provided a goal for research by providing Earth-life parameters to use in searching for habitable exoplanets. Participants suggested increased interdisciplinary research, especially with astrobiologists, as good practice to manage the N=1 Problem moving forward in the search for habitable exoplanets. Results should be seen in light of the small sample size of research participants, but can be read as a set of impacts the N=1 Problem *can* have on the exoplanet science community. Only having one example of a planet with life affects how humans search for exoplanets, just as it affects how SF creators present exoplanets in science fiction. This study provides valuable social science data on how researchers manage unknowns in science.

Chapter 8 concludes with a summary of my findings and a discussion of potential applications of my research. I have used Bayesian network methodology to quantitatively study the portrayal of exoplanets in science fiction, finding that SF does respond to new scientific discov-

eries. This evidence of the reciprocal relationship between science and SF suggests that SF has a strong potential to be used in science education work, perhaps in introducing concepts and providing an inspirational jumping-off point from which educators can help students conduct their own research and place science into context, as suggested by the results of the StACES anthology project. The N=1 Problem study highlights challenges for scientific research in a field with lots of unknowns, like the search for habitable exoplanets. In addition to the strategies interview participants discussed, the applied speculation of science fiction is another way to investigate the unknowns of exoplanet science. I argue that science fiction is valuable not only as art but as a tool for science, providing a human perspective on and exploring ramifications of new scientific breakthroughs. SF provides a shared reference point for many, and its use of narrative lends it to educative pursuits; future work could include expanding my Bayesian network results to include finer categories and using the results to inform lesson plans and other science education tools. My research uses a combination of quantitative and qualitative methods studying the portrayal of exoplanets in science fiction to show that changes in science effect SF, and that SF can be used for science education.

<div style="text-align: right">

1

</div>

Exoplanet science

"Bequeathing worlds till thought's vast cope o'erruns
Thine eyes, soon after birth,
Relinquish'd thine own earth–
Unwearied watcher of a thousand suns!"

—R.H. Horne, *Sir John Herschel*

[Portions of this chapter are taken from my work in the paper "Science Fiction Media Representations of Exoplanets: Portrayals of Changing Astronomical Discoveries" published in the March 2024 issue of JCOM Journal of Science Communication.]

The discovery of exoplanets in the 1990s forever changed the field of astronomy by extending our understanding of planets beyond the solar system. The effects of this discovery are ongoing, and in March 2022, the NASA Exoplanet Archive announced 5000 exoplanets had been confirmed by its criteria, an arbitrary milestone which is nevertheless emblematic of the growth of the field (Brennan, 2022). The intervening three decades have seen the rise of exoplanet science, a new sub-field of astronomy dedicated to detecting and characterising these

distant planets as well as inferring their demographics. Exoplanet science has become a research and funding priority, with major space agencies like the European Space Agency (ESA) and the National Aeronautics and Space Administration (NASA) highlighting it in their long-term planning publications (ESA Voyage 2050, 2021), (National Academies Decadal Survey, 2021). Although over 5000 exoplanets have been discovered, the usually indirect detection methods used result in most exoplanets having few constrained parameters. Exoplanet science missions provide pieces of data about exoplanets, like mass or radius, and scientists fill in the gaps with modelling based on fundamental physics and chemistry. In this chapter, I provide an overview of the portions of exoplanet science relevant to this thesis, and introduce science fiction (SF) as a genre of speculative media that frequently features exoplanets and employs an imaginative form of world-making to construct them. SF stories, though fictional, have a basis in scientific fact, and SF stories of exoplanets are attempts to fill in gaps in the data—imaginatively. Presenting the findings of exoplanet science shows why this field of astronomy, by virtue of the recent, impactful discovery of exoplanets, is an ideal one with which to investigate the relationship between science and SF.

1.1   Early concepts of exoplanets

Cosmic pluralism is the idea that there are other worlds (i.e. planets) out there that may be like our own. The concept dates back at least to the Ancient Greek atomists (Dick (2020c); Epicurus (1925)). The plurality of worlds is often tied both historically and in the modern literature to the concept of extra-terrestrial life, given it is a natural leap from the idea of other planets to wondering whether anyone might live there. While it is beyond the scope of this thesis to delve into comparative international historical cosmologies, it is important to mention that the majority of the scholarship on cosmic pluralism that I cite comes from the Western tradition, that is, the line of work stretching from the Greeks to medieval Middle Eastern cultures to Copernicus and beyond. Some scholars, like Dick, who claims that cosmic pluralism is unique to traditions of Western thought and wonders why this might be (Dick, 2020a), and Krupp, who goes further and states "there is no room in the indigenous cosmologies of the Americas for the plurality of worlds and extraterrestrial life" (Krupp, 2015), consider cosmic pluralism a Western idea. I note that there is currently increased research into indigenous cosmologies, and refer interested readers to (Szerszynski, 2019). Many cultures claim they come from specific stars, and people from other worlds feature in non-Western precursors to

science fiction (Gillon, 2019) (more on this in Chapter 2). However, the crucial takeaway is that all of these cosmologies—those that do and do not allow for multiple worlds—remained in the realm of philosophy until the Copernican revolution.

In the 16th century, Polish astronomer Nicolaus Copernicus's publications began the shift in astronomical understanding from a geocentric to a heliocentric—Sun-centred—solar system (Copernicus (1543)). Copernicus's finding finally lent empirical evidence to cosmic pluralism, by indicating the Sun might be like other stars. As Jaime Green argues, our imaginations are good, but they can only extrapolate so far from what is known—they need to be prompted, and Copernicus provided that prompting (Green, 2023). Copernicus "gave birth to a new tradition where the term world (mundus) was now redefined to be an Earth-like planet, and each of these Earthlike planets took on the kinematic or motion-related functions of the single Earth in the old geocentric system" (Dick, 2020c). Early adopters included Giordano Bruno, who extended Copernicus's idea to other stars, which he proposed might be orbited by exoplanets—though he did not yet use that term (Gillon, 2019). While Copernicus lent the model, Bradley in the 1700s proved the theory by measuring the aberration of light (Bradley (1727)). With the idea that other stars were like our Sun, and might host planetary systems of their own, the seeds for the eventual scientific search for exoplanets were planted.

As is often the case, theory preceded observation. Maunder's *Are the planets inhabited?* ("round our Sun there is but a narrow zone in which a habitable world may circle") and Shapley's "liquid water belt" concept ("the living planet must be at a proper distance from its star—in the liquid water belt—not as close as Mercury is to the Sun nor as remote as Jupiter") both evolved into the modern concept of a *habitable zone* and anticipated observational evidence of exoplanets by 82 and 42 years, respectively (Maunder, 1913); (Shapley, 1953); (Kasting et al., 1993). These writings provide another example of the link between searching for other worlds, and scientific considerations of extra-terrestrial life. For more on the concept of the habitable zone, see Chapter 4 and Chapter 7.

1.2 First detections

Detecting planets orbiting other stars presents an observational challenge. Planets are both dim and located close to much brighter stars such that the most straightforward detection method, directly imaging planets, is difficult (it was not accomplished until 2004 (Chauvin,

G. et al., 2004)). Before the first successful detection, there were many attempts and several false detections of exoplanets. Already in the 1690s, the Copernican revolution had spurred astronomer Christiaan Huygens to conceive of detecting exoplanets, but he concluded the technology did not yet exist (Huygens, 1698). In the 1960s, Peter van de Kamp at Swarthmore College made several claims to have detected Jupiter-sized planets orbiting Barnard's Star using the astrometry method. This method painstakingly compares multiple images of the same star to detect motion due to a gravitational wobble indicative of an orbiting body (van de Kamp, 1969). Van de Kamp's discoveries were ultimately found to be spurious due to instrument error, as they corresponded to times the lens had been cleaned (Tasker, 2017). In the 1980s, several possible detections were made using the radial velocity method (explained in detail later), later to be successful in finding 51 Pegasi b, but uncertainties in either the stellar mass or planetary mass made these finds impossible to classify as certain planets beyond scientific reproach (Latham et al., 1989); (Campbell et al., 1988). Bruce Walker, who was involved in these pre-51-Pegasi radial velocity surveys, later reflected: "It is quite hard nowadays to realise the atmosphere of skepticism and indifference in the 1980s to proposed searches for extra-solar planets. Some people felt that such an undertaking was not even a legitimate part of astronomy" (Walker, 2012). While exoplanets remained the realm of science fiction, these early planet-finding attempts helped build the level of instrument sensitivity necessary to achieve eventual success.

Although 1995 is cited throughout this thesis as the year of the first detection of exoplanets, there were earlier discoveries made in a very unexpected environment. In the late 1980s and early 1990s, a team found planets orbiting a pulsar, a remnant of a massive star left over after the star has exploded in a supernova. These planets, one of which has only twice the mass of the Earth's moon (thus making the least massive exoplanet known one of the first discovered), were found using a different methodology called pulsar timing (Wolszczan and Frail, 1992); (Wolszczan, 1994). Pulsars have radio jets that rotate extremely precisely, and minute variations in the timing of this rotation can indicate the presence of a planetary system. This makes the pulsar planets find more unique both in stellar environment and in detection method—this method cannot be used to find planets around main-sequence stars, that is, those fusing hydrogen to helium in their cores for a long, stable time span that may enable the emergence of life. Amid many questions on how the planets formed in the first place, the pulsar planets did not represent the same momentous proof of exoplanets around main

sequence stars that the 1995 discovery did. This is the reason the clarification "first exoplanet discovered around a sun-like star" is appended to the 51 Pegasi b discovery. 1995 is widely accepted in the scientific community as the year exoplanets were discovered, with the 51 Pegasi b discoverers winning the Nobel Prize for their effort.
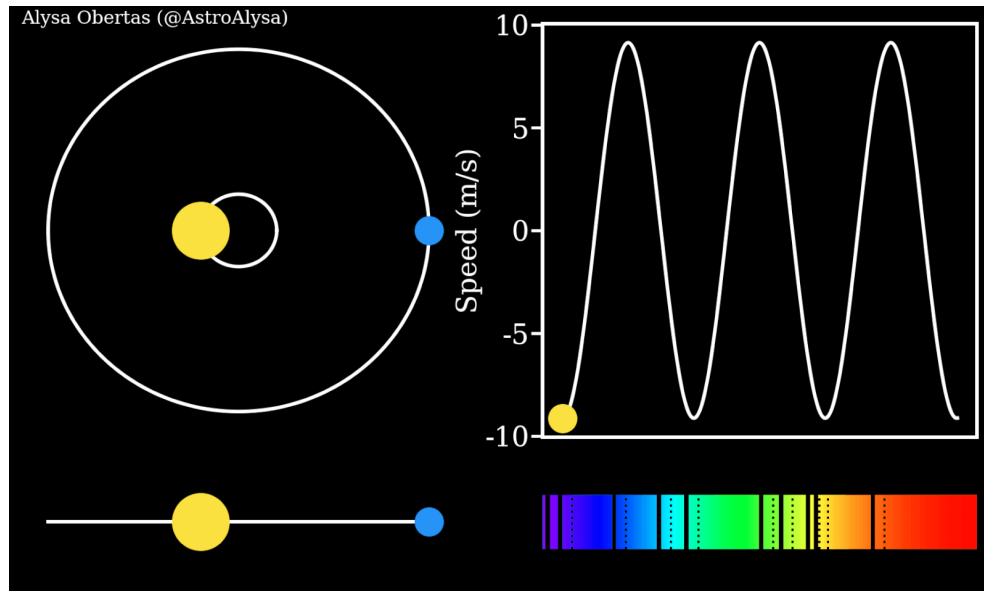
Figure 1.1: A simplified graphic of the radial velocity of exoplanet detection. This image was originally a moving gif; it is paused when the spectrum is most blue-shifted. Note how the dark lines are much bluer than their corresponding dotted lines, at the expected location of the line. (Credit: Alysa Obertas, CC BY-SA 4.0, https://creativecommons.org/share-your-work/licensing-considerations/compatible-licenses/)

In October 1995, Michel Mayor and Didier Queloz announced the discovery of 51 Pegasi b, a planet more massive than Jupiter and orbiting around its Sun-like (main sequence) star in a blistering 4.2 days (Mayor and Queloz, 1995). For comparison, Mercury, the closest planet to the Sun in our solar system, takes 88 days to complete one orbit. The long time gap between theory and observation was due to the difficulty of exoplanet detection; the first successful detections utilised indirect methods through observations of their host stars. Mayor and Queloz used the radial velocity method (see Figure 1.1). This method takes advantage of the fact that a star and its planetary system orbit around their joint centre of mass, which is displaced from the centre of the star. Thus, the star appears to "wobble". By measuring the star's spectrum over time and looking for Doppler shifts, or shifts towards bluer or redder wavelengths depending if the star is moving toward or away from our line-of-sight, this wobble can be detected and measured. The star 51 Pegasi moves about 60 metres per second due to its planetary companion; for comparison, an Earth-mass planet orbiting at 1 astronomical unit (AU),

or Earth-Sun distance, would cause a radial velocity signal of just 9 centimeters per second (Lovis and Fischer, 2010). The precision required to detect such a planet is still beyond our current instruments, but future space missions like PLATO (PLAnetary Transits and Oscillations of stars) will be able to do so (Rauer et al., 2014).

1.3 Increased discoveries and characterisation

A few years after the first radial velocity discoveries, a new method leapt onto the scene: the transit method. The transit method is simpler to explain conceptually: when a planet passes between its star and an observer, the light the observer sees from that star dims slightly. From the decrease in stellar flux, the radius of the planet can be determined, and from the period between successive decreases, the orbital period can be found (where a year is the Earth's orbital period) (Fischer et al., 2014). See Figure 1.2 for a visualization. The transit method was first used on HD 209458b, a planet which had already been discovered using the radial velocity method (Charbonneau et al., 1999). At the time of writing, 75% of exoplanet detections have been made via the transit method (NASA Exoplanet Archive, 2024).
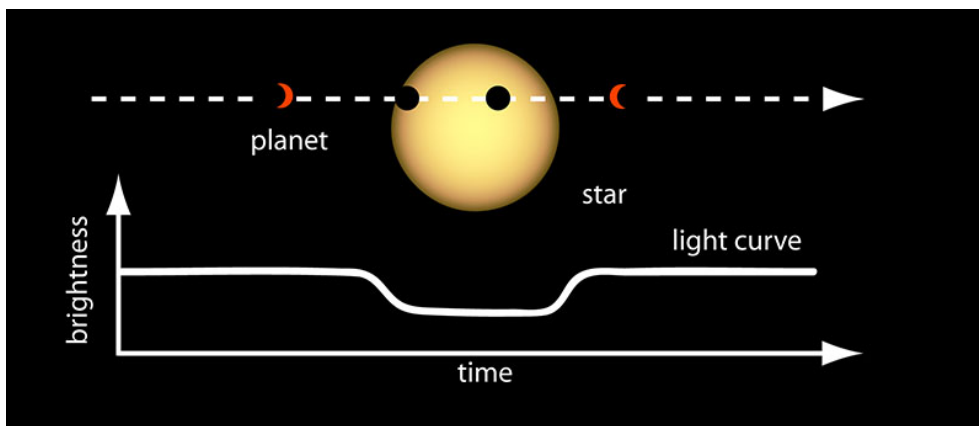
Figure 1.2: A simplified graphic of the transit method of exoplanet detection. As the planet passes between the observer (on Earth) and its star, the flux the observer sees from the star decreases. The graphic is not to scale; actual decreases tend to be <2% of the stellar flux. (Credit: NASA Ames)

From the first detection, the discovery of exoplanets has completely changed models of solar system formation and planet categories. 51 Pegasi b is a massive, gaseous planet orbiting much closer to its star than Mercury does the Sun. At the time of its discovery, this was completely unexpected: it had been thought that a solar system architecture like our own, with small, rocky planets close to the star, and massive, gaseous planets further from the star, was standard. 51 Pegasi b belonged to an entirely new class of planets termed "hot Jupiters". The

formation of such planets is an ongoing area of study: it is debated whether they formed in situ, or migrated inward to their present positions from a more distant original orbit (Dawson and Johnson, 2018). Other new classes of planets have been discovered, such as Super-Earths and Mini-Neptunes, those planets with masses intermediary between Earth, the largest terrestrial planet in our solar system, and Neptune, the smallest gas giant. A statistical "radius valley" has been found in which planets in the 2-Earth-radius regime are rarer; why this is provides another active area of research in planet formation and evolution (Venturini, Julia et al., 2020). Planets have also been found with different types of stars than the Sun, including in red dwarf systems (M dwarfs). Red dwarfs are the most common type of star and are smaller and cooler than the Sun. To orbit a red dwarf in the habitable zone, a planet is often close enough that gravitational tidal forces have slowed its rotation such that one hemisphere permanently faces towards the star and the other permanently faces away; this condition is called being "tidally locked". Tidal locking is not unique to planets around M dwarfs, but to close-in orbiting bodies: our moon provides one such close-to-home example of this phenomenon, being tidally locked to Earth. Given the tidally locked environment and the increased flare and XUV flux of red dwarf stars, potential habitability of planets in such systems forms another active area of research (Barnes et al., 2009); (Sagear and Ballard, 2023). These exoplanets have no analogue in our solar system and are therefore of great scientific interest. It is only in the past few decades that astronomers have been able to compare our own solar system to others, and begin to learn where we might be unique and where we might be standard. The discovery that our solar system may not be "average" can be considered a continuation of a centuries-long de-centring of humanity that began with the Copernican revolution.

Depending on the detection method, astronomers can determine mass, radius, orbital eccentricity (how circular or elliptical an orbit is), inclination, and orbital period (via semi-major axis) of exoplanets. Given mass and radius both, a density can be calculated, giving insight into what materials might make up the planet. Some of these physical properties come with wide error bars. Only a minimum mass estimate can be calculated for a planet detected via radial velocity, as the inclination (the angle or tilt at which we see it) of its solar system is unknown (Lovis and Fischer, 2010). Astronomers can also identify molecules in atmospheres of transiting exoplanets through the spectra (more in the next paragraph). Ingenious detection methods have allowed astronomers to find exoplanets we cannot even see directly, yet there remain many unknowns about each discovery. Detection methods can also be biased

towards planets that are easier to detect: hot Jupiters, as massive planets with short orbits, have large gravitational effects on their stars and block out bigger percentages of stellar flux when transiting, therefore they are easier to detect via the radial velocity and transit methods. Large planets on wide orbits, away from overwhelming stellar glare, are easier to directly image, and so directly imaged planets often match these characteristics. Smaller, rocky planets are harder to detect by all current methods—indeed, it was not until 2009 that the French CoRoT (Convection, Rotation et Transits planétaires) mission discovered the first rocky exoplanet (Léger, A. et al., 2009), 14 years after the discovery of 51 Pegasi b. The mass-period distribution, colour-coded by detection method, is shown in Figure 1.3 and demonstrates the current discovered exoplanet population. The chart also includes discoveries from other methods of exoplanet detection including microlensing, which detects planetary masses based on the gravitational lens effect, timing variations, which utilises variations in timings of transits, orbital brightness modulation, which measures starlight reflected by exoplanets, and astrometry, described previously in section 1.1. Given the youth of the field of exoplanet science, we are only just beginning to perform population statistics on discovered worlds and extrapolate about the overall exoplanet population, of which the 5000-odd worlds discovered constitute a tiny fraction.

Exoplanet science is broadly moving from discovering planets to characterising them, including learning more about exoplanetary environments. This shift towards characterisation is reflected in current and upcoming exoplanet missions. While the trailblazing Kepler mission's goal was to find out how common planets are by observing one small area of the sky (Borucki et al., 2010), upcoming decades will see the implementation of ELTs (Extremely Large Telescopes) with capabilities to determine the chemical makeups of exoplanetary atmospheres, focusing on bright host stars to capture the necessary detailed spectra (Kaltenegger, 2017). The James Webb Space Telescope, launched in late 2021, uses transmission spectroscopy, comparing spectra of exoplanet host stars when the exoplanet is and is not transiting to determine which chemical species are present in the exoplanet's atmosphere (Beichman and Green, 2018). Early JWST data already provided evidence for water and clouds in the atmosphere of exoplanet WASP-96 b by summer 2022 ("NASA's Webb", 2022). The European Space Agency's planned upcoming PLATO and ARIEL space telescope missions will constrain masses, radii, densities, stellar irradiation, and ages of exoplanetary systems, and observe thermal structures and atmospheric compositions, respectively—all crucial pieces of information for assess-
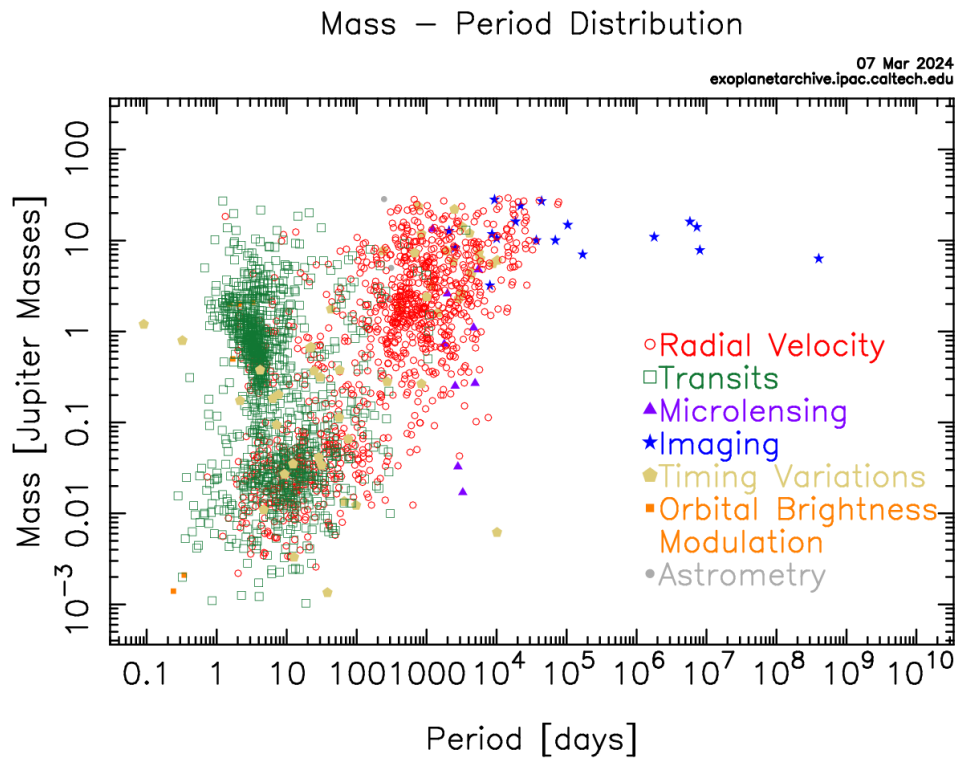
Figure 1.3: The mass-period distribution of all discovered exoplanets. Most exoplanets found thus far are more massive and have shorter years than the Earth. (Credit: NASA/JPL-Caltech)

ing exoplanet habitability (Rauer et al., 2014); (Tinetti et al., 2018). PLATO will specifically look to detect terrestrial planets in the habitable zone, aiming for a precision of 3% in radius measurements (PLATO Definition Study Report, 2017). These missions will bring this young field experiencing explosive growth closer to understanding exoplanets as whole places with unique environments.

1.4 Two types of modelling

In just a few decades, humanity has gone from having only our solar system as examples of planets, to thousands. As observational astronomers work to learn more about each exoplanet, theoretical astronomers use computer modelling to anticipate exoplanetary characteristics: for example, what global atmospheric dynamics and evolution might be like in various stellar and planetary environments (Bonfanti, A. et al., 2021), and how solar systems form from protoplanetary discs (Bitsch, Bertram et al., 2015). This type of scientific modelling is based on understandings of laws of physics, and often on available observational data. Lisa Messeri, an anthropologist who did a study on exoplanet scientists, argues that they employ a process of imagination to picture a world entire from a graph, such as a transit light curve, or the

results of a computer model. Given that pictures often do not exist of exoplanets (certainly not pictures crisper than a few pixels), observational and theoretical exoplanet astronomers alike must engage in an act of place-making from data and figures very divorced from what one typically imagines when one thinks of a planet. "One who becomes a successful exoplanet astronomer is one who becomes a successful crafter of worlds" Messeri writes (Messeri, 2016). Exoplanet scientists must understand a planet from the data, and Messeri argues this requires some use of imagination even amongst the scientific method. Messeri's work evokes Latour's concept of "circulating reference", in which a chain of transformations separates real natural phenomena from their presentation in scientific diagrams (Latour, 1999). Literature scholar Will Tattersdill writes of Latour "his model can in some respects work for fiction as well; that the strongest fiction is not that which is most dissociated from its opposite, fact, but that which is able to establish a retraceable chain of reference to it, whatever the length of that chain" (Tattersdill, 2016). As will be discussed in chapter 2, SF does involve inspiration from and basis in scientific fact, and I will continue Tattersdill's extension of the chain of reference to fiction.

I argue this process of 'crafting worlds' among exoplanet scientists, of envisioning a planet from a graph or of using a computer model to imagine an exoplanetary atmosphere, is, despite a difference in methodology between sciences and humanities, not altogether unlike the process of worldbuilding in science fiction. Both computer modelling of exoplanets, and worldbuilding of fictional exoplanets in science fiction, are ways of filling in the gaps in our current knowledge. SF represents another body of writing in which people have been imagining exoplanets, and for hundreds of years longer than they have been confirmed observationally. Chapters 4 and 5 of this thesis will present a database of science fiction exoplanets and the results of a Bayesian network analysis of said database, with the aim of investigating the interplay between science and SF. Chapter 6 presents a research project on how teams of scientists and writers work together to include science in SF worldbuilding, and Chapter 7 addresses how the "N=1 Problem", the fact that we only know of one planet with life, affects human understandings, both imaginative and scientific, of concepts of other life in the universe. Exoplanets form a useful dataset with which to investigate the portrayal of science in SF due to their momentous discovery, and long-term presence in SF. This study is of interest to science communicators, as well, given that SF exoplanets represent much of the public's exposure to exoplanets: how is science being presented in the fiction? Astronomer Jessie Christensen, in

a press release about the 5000-exoplanet milestone, said "Each one of them is a new world, a brand-new planet. I get excited about every one because we don't know anything about them"(Brennan, 2022). Christensen's statement "we don't know anything about them" is, of course, objectively incorrect in that we do know certain parameters as explained in this chapter—but she does not mean the statement to be objective. Instead, she is referring to the unknown, the gaps to be filled as planet after planet is discovered. The concept of exoplanets is one that scientists and fiction writers have both been theorising and imagining for a very long time, but that we only recently have evidence for.

2

Science fiction and science: a reciprocal relationship

"Science does not know its debt to imagination."

—Ralph Waldo Emerson

"My fondest hope for this book is that it will be made obsolete by the pace of real scientific discovery."

—Carl Sagan, *Author's Note, Contact*

[Portions of this chapter are taken from my work in the paper 'Science Fiction Media Representations of Exoplanets: Portrayals of Changing Astronomical Discoveries' published in the March 2024 issue of JCOM Journal of Science Communication, my paper 'Dialogues Between Science and Fiction in the Creation of an Anthology' published in BSFA Vector issue 297 (2023), and

my chapter 'Through a Mirror, Earthly: Solaris, Gaia, and the Search for Other Worlds' in the collection Stanisław Lem and His Aliens: A Tribute and a Challenge (2022)]

This chapter contextualises my research by situating it within a tradition of the scientific study of literature, from the work of the Russian Formalists to the field of digital humanities today. It will introduce science fiction as a genre, focusing on its reciprocal relationship with science—the mutual influences between the two, and how neither is static but instead both evolve together. The influence of science on SF gives the genre the potential for use in teaching science. In order to investigate this, it is necessary to study SF's overall portrayal of science using larger-scale data science techniques, rather than close reading or case studies—precisely what my Bayesian network project aims to do. This goal motivates my research methodology.

SF is a significant transnational cultural phenomenon. SF films consistently take 10-20% of the North American box office market share (Nash, 2023). SF books like Cixin Liu's *The Three-Body Problem* and Andy Weir's *The Martian*, and video game franchises like *Mass Effect* sell millions of copies worldwide (McGregor, 2020); (Alter, 2017); (Radić, 2021). SF is not merely a literary genre—SF fans read, watch, listen to, and play SF stories, and this is reflected in the source works of the fictional exoplanets in this study (see Chapter 4). Given this popularity, there is growing interest in analysing SF's utility for science education (Irani and Weitkamp, 2023), and the role SF plays in imagining futures for scientists and the public alike (Reinsborough, 2017).

2.1 Definitions and origins

It is well-agreed that most fiction set on exoplanets fits under the generic definition of SF—such stories are far from 'edge cases' of the genre—but I begin with a discussion of the definition of the SF genre to emphasise SF's unique relationship with empiricism in its worldbuilding. *Worldbuilding* is the process by which authors construct a fictional world from the ground up, including its environment, social aspects, economic aspects, and more. I argue SF's *cognition effect* (see next) leads to a distinctive relationship between science and SF writers, which is not found in other genres.

Much ink has been spilled on the topic of defining science fiction; for purposes of this thesis, I focus on what various definitions say about SF's relationship with science. Compared to other types of genre fiction (such as romance, detective fiction, or fantasy), science fiction

defies simple categorisation. A narrower definition, for example, limiting SF to stories set in space, would exclude Aldous Huxley's *Brave New World* or Mary Shelley's *Frankenstein*, as well as any number of stories featuring robots or artificial intelligences. On the other hand, painting too broad a definition risks including nearly anything with a speculative element (Freedman, 2000). Some scholars, like Slusser, emphasise common themes of the genre, like that it is "future-directed" or "open-ended" without laying out a specific definition (Slusser, 2003). My readings are founded on the theorist Darko Suvin's more operative definition, that SF relies on "estrangement and cognition" and features an "imaginative framework alternative to the author's empirical environment" (Suvin, 1979). In other words, this means that SF features at least one significant change (estrangement) from reality, which is presented cognitively in a way that distinguishes it from fantasy—SF works must account for their worlds rationally within the text. The 'cognition' is the science that separates the genre from fantasy; while fantasy universes operate on the supernatural, science fiction universes operate on natural laws (within the story). In the worldbuilding of the science fictional text, the estrangements are treated as science whether or not they are consistent with real-world science, rather than being left to magic and mystery. For example, in fantasy *Harry Potter*, characters move faster than light by using magic to teleport, whereas in science fiction *Star Trek*, characters move faster than light by using the warp drive, the in-universe design of which references the real science of matter-antimatter reactions and the Lorentz factor from special relativity to achieve a fictional effect. The *estrangement* is travelling faster than light, and *Star Trek* presents it cognitively while *Harry Potter* does not. Some SF authors emphasize the genre's strong relationship with science in their own definitions, with Asimov defining SF as "literature which deals with the reaction of human beings to changes in science and technology" (Asimov, 1975). There is a broad range of scientific accuracy in SF; the term 'soft' refers to SF with less basis in scientific fact (for example *Star Wars*), and the term 'hard' refers to SF with more basis in scientific fact (for example *The Martian*). SF, however "soft", takes its estrangements from science-based ideas. I propose that there is always a basis in an idea from science, upon which the author then extrapolates, and it is this cognitive estrangement of real science that allows for creative extrapolation of and speculation on real-world science in science fiction.

In addition to generic definitions, scholars debate which work represents the earliest example of science fiction. The term 'science fiction' was not popularised until the early 20th century—in the 1920s Hugo Gernsback, founder of the first science fiction magazine *Amazing Sto-*

ries, and the man for whom the Hugo Award is named, was calling it "scientifiction"—though of course scholars have looked back for examples of the form that predate the term (Stableford, 2003). While much of the exploration of scientific inventions in Western fiction naturally comes during and after the Scientific Revolution of the 16th and 17th centuries (of which the Copernican revolution was part) (Stableford, 2003), even before the Copernican revolution, there were several early works that bear some hallmarks of the eventual genre, and as indicated in the previous chapter they come from different parts of the world. Lucian's *A True Story* (which begins by saying that none of its contents are true), written in the 2nd century CE in Syria, features alien life on the sun and moon. In addition to these solar system peoples, Lucian has "the Cynobalanians, that were sent from the Dogstar to aid him: these were men with dogs' faces, riding upon winged acorns", hailing from Sirius ("the Dogstar") (Lucian, 1902). *The Tale of the Bamboo Cutter*, a 10th-century Japanese story, features a princess from the moon (Sanjin, 1908). "The Adventure of Bulukiya", part of the Islamic Golden Age epic *One Thousand and One Nights*, has some science fictional elements in the protagonist's visit to other worlds—although these worlds are across a vast mythic sea, rather than in outer space (Burton, 1885). The novel *Somnium* was even written by a scientist, Johannes Kepler, in 1608, and features life on other worlds such as the moon (Kepler, 1634). But while all these examples constitute "imaginary voyages", a literary subgenre, they do not all feature humanity engaging with scientific changes the way true science fiction does, as Brian Aldiss argues (Aldiss, 1973). The inspiration from science is missing.

For many works, take for example *A True Story*, it is impossible to tell whether Lucian is referring to people coming from the Dogstar in the sense of extra-terrestrials from another planet, or more akin to the ways cultures have mapped their mythology onto the heavens for thousands of years (e.g. the mythological hunter Orion in a constellation). Simply going to a destination in outer space is not enough to render a work science fiction—though works set in outer space do almost always feature the cognitive estrangement of interacting with, speculating on, and exploring the ramifications of scientific ideas. Aldiss considers the first example of science fiction to be Mary Shelley's *Frankenstein* in 1818, in which the story features the protagonist engaging with discoveries in electricity and galvanism (Aldiss, 1973). The fact that some other literature scholars, like Stableford, see *Frankenstein* as "anti-science" fiction because it offers a critique of scientific ethics rather than presenting all progress as morally good, is a misunderstanding of science fiction (Stableford, 2003). SF has always presented

human interactions with science, and in fact has a long history of featuring ethical and moral challenges that arise with scientific discovery. By the turn of the 20th century, writers like Wells and Verne were further establishing the genre, and by the mid-20th-century, science fiction was well-known in the Western fiction market (Stableford, 2003).

There are a few examples of exoplanets in science fiction before the 20th century. The titular character of Voltaire's novella *Micromégas* is a giant from a planet orbiting the star Sirius—employing the Copernican revolution knowledge that other stars might have planetary systems like our sun does (Voltaire, 1752). Longer ago still is Margaret Cavendish's *Blazing-World*, written in the 1660s, in which the main character is a young woman who travels from the north pole of the Earth to another planet orbiting another star. Cavendish's work seems to be postulating our sun might actually be in a binary system, with a companion star we cannot see because our sun is bright and blocks it from view. Cavendish's narrator explains: "You must know, that each of these Worlds having its own Sun to enlighten it, they move each one in their peculiar Circles; which motion is so just and exact, that neither can hinder or obstruct the other; for they do not exceed their Tropicks: and although they should meet, yet we in this World cannot so well perceive them, by reason of the brightness of our Sun, which being nearer to us, obstructs the splendor of the Sun of the other World, they being too far off to be discerned by our optick perception, except we use very good Telescopes; by which, skilful Astronomers have often observed two or three Suns at once" (Cavendish, 1668). Here Cavendish references the use of telescopes in astronomy to observe "two or three Suns at once", showing she had read about astronomical discoveries of binary and multiple star systems. Cavendish, the first woman to attend a meeting of the Royal Society, was well-read on current science and wrote non-fiction books on scientific topics as well, and these influences are clear throughout *Blazing-World* (Holmes, 2010). In works before Voltaire and Cavendish, like those from Lucian, *Bamboo Cutter*, and *Thousand and One Nights*, it becomes very difficult to differentiate fictional exoplanets from portrayals of alternate dimensions or realities. Cavendish's is the earliest fictional example I have uncovered that is clearly a planet orbiting another star.

Exoplanets did not become frequent settings for science fiction stories until the mid-20th century. As Stableford writes "the gradual removal of terra incognita from maps of the Earth's surface helped to force utopian and satirical images out into space" (Stableford, 2003). With most land on Earth explored, including the north and south poles by the 1920s, stories set on

Mars and Venus, like Edgar Rice Burroughs' *A Princess of Mars* (1912) and Robert Heinlein's *Space Cadet* (1948), became popular as imaginations took to outer space. Later, as Gillon writes, "the revelation that Mars was a dead inhospitable world combined with the understanding of the structure of the Milky Way made the imagination of science fiction (Sci-fi) writers turn to the stars" (Stableford, 2003); (Gillon, 2019). Mars was understood to be "a dead inhospitable world" after the first flyby missions in the 1960s. These examples show SF's subjects and settings changing in response to scientific discovery—if a writer wants to have a reasonably-scientifically-plausible work and feature intelligent life on another planet in it, they must set that work outside our solar system. Writers are indeed very interested in exploring extra-terrestrial life (see Chapters 4 and 7), and after a spate of missions in the 1960s and 70s showed both Venus and Mars as clearly inhospitable to Earth-like life, our neighbour planets were removed as plausible settings. Instead, exoplanets became the setting to use for SF stories in which writers wanted intelligent extra-terrestrials, and indeed, most SF exoplanets in my database (see Chapter 4) come from the mid-century or after.

2.2 Mutual influence of science and SF

Beyond exoplanet SF, SF and science in general have a unique, reciprocal relationship. Many scientists write science fiction—Isaac Asimov and E.E. Smith for example—and many SF authors are avid supporters of science programmes and science communication (Stepney, 2015). Creators of SF literature and film and television often refer to science consultants for accuracy, and workshops like the NASA-funded Launchpad, which aimed to teach writers about science for their books, are not uncommon—the Hugo-award winning author N.K. Jemisin was inspired to write the *Broken Earth* trilogy after interacting with scientists at a Launchpad workshop (Khatchadourian, 2020). Often, much focus is given to the sources and accuracy of the science that creators present in their SF, as evidenced by the convention of SF authors often referencing their research in the acknowledgements sections of their works.

The opposite question—why is SF important for science?—is asked more rarely. In order to understand the relationship between SF creators and scientists, it needs to be asked. In her acceptance speech for her Hugo, Jemisin addressed this question by describing SF as "the aspirational drive of the Zeitgeist", with creators as "the engineers of possibility" (Jemisin, 2018). I asked Mary Robinette Kowal why science fiction was important for science at a Caltech-hosted panel on SF worldbuilding, and she answered "we can [...] ask these what-if questions, and

then attach it to the way it affects people [. . .] we can model for you" (Kowal, 2020). Kowal and Jemisin describe SF creators as "engineers" and use scientific terms like "modelling", showing that creators extrapolate on the present. Scientific achievements do not happen in a vacuum—they always have human effects, and SF examines what these effects might be. For example, while a scientist might run a computer model to determine if a planet's orbit could be stable in a binary star system, an SF writer might "run a model" of how life in such a system, with two suns, would affect a society that lives there or a human that travels there. In addition to providing such explorations of possible futures, SF inspires countless scientists in their career paths. A survey of 239 UK astronomers attending the 2022 UK National Astronomy Meeting provided statistical evidence that the majority were interested in SF and stated that SF influenced their career decisions (Stanway, 2022). One aim of my research is to tease out how this mutual relationship works between SF and scientists.

Lem's *Solaris* provides an excellent example of how SF can critique and comment on science, through imagining a fictional field of science and revealing its inadequacies. The story models a situation of first contact with an intelligent extra-terrestrial—the Solaran ocean—and exploration of an exoplanet. The human scientists in the story have a tendency to anthropomorphise and adhere rigidly to Earthly categorisation methods, ultimately failing to communicate with the ocean. "We don't need other worlds. We need mirrors" as one character puts it, describing the trouble as being caused by humanity taking to the stars while not having taken the time to fully understand their own world first (Lem, 2014). As Lem's 1961 book was diagnosing a current problem in science, that same problem was being addressed by emerging ideas. These included the Overview Effect—the effect of humanity seeing the Earth as a whole system for the first time during the Space Age—and Gaia theory, the idea that the Earth and its lifeforms are one huge interconnected system, which evolved together and self-regulates to perpetuate conditions for life. Lem's planet Solaris is fundamentally a Gaian planet, as evidenced by the Earth scientists' attention first being drawn to it by its self-regulating of its orbit to maintain habitability. The same scientists initially do not recognise the planet as a living being due to conflicting methodologies and definitions among their biologists and physicists. Today's scientists are no longer quite so siloed in their fields as Lem's fictional ones, and the ultimate lesson of the book, which advocates for interdisciplinarity and self-awareness in the search for other worlds, has been to an extent taken to heart by astrobiologists and exoplanet scientists who frequently use Earth as an analogue for exoplanets and focus on understanding

Earth in preparation for learning more about other worlds (Puranen, 2022).

This relationship between SF and science often leads to a common misconception that science fiction predicts the future. The similarity between the communicators on the original 1960s Star Trek series and early flip-phones is often cited as an example (Venables, 2013). The answer is more complicated yet less satisfying—it is that each inspired the other, or as Le Guin famously put it "Science fiction is not predictive, it is descriptive" (Le Guin, 1969). Authors tap into the feelings of the time and develop what-if scenarios. If these scenarios occasionally come true, it is because the seed was already there. Science fiction, for all its alien vistas and strange biology and travels to distant galaxies, is always a commentary on the human condition and the writer's contemporary world, and science and SF are a revolving door. I suggest it may be analogous to Latour's concept of "circulating reference", in which a chain of transformations separates real natural phenomena from their presentation in scientific diagrams (Latour, 1999). Similarly, a chain of transformations might separate real natural phenomena from their presentations in science fiction. Uncovering this circulating reference in science fiction worldbuilding to understand the portrayal of science in SF constitutes my contribution to the study of the relationship between science and SF.

To explain this, I employ as an example the discovery of the TRAPPIST-1 system of exoplanets, which was published in Nature in 2016 (Gillon et al., 2017). The TRAPPIST-1 system contains seven terrestrial exoplanets in tightly-packed orbits around a red dwarf star 39 light years from Earth. The discovery garnered significant attention due to the presence of multiple of these planets potentially in the liquid-water habitable zone (Turbet et al., 2020). A short science fiction story by Suhner was published in *Nature* alongside the main article, featuring a future human living on the planet TRAPPIST-1e, which is tidally locked so the same side is always facing the star (Suhner, 2017). Even before publishing their science results, the discoverers of the TRAPPIST-1 system had clearly discussed with an SF author to create a fictional portrayal of their findings (for more on science consultants for science fiction, see Chapter 6). This alone demonstrates the inter-connectivity of science and SF, but soon after the discovery was made public, players of a space-faring video game called *Elite Dangerous* noticed that there was already a solar system in the video game of seven terrestrial planets 39 light years away. *Elite Dangerous* uses an algorithm that, through procedural generation, generates a 1:1 scale Milky Way galaxy for its players to traverse. The solar systems are generated through actual scientific data about solar system formation, stellar age, temperature, and metallicity,

mass distribution in systems, etc. Given the presence of this system, players claimed that *Elite Dangerous* had "predicted" the TRAPPIST-1 system (Good, 2017). The system in the game was not originally an exact match for TRAPPIST-1, as it orbited a brown dwarf instead of a red dwarf star, but the game creators have since altered the parameters of their existing system so that it now matches what we know of the real TRAPPIST-1 system (Good, 2017). This example illustrates how SF can extrapolate based on known science, then science sometimes proves those extrapolations, and then science fiction sometimes alters itself to match. SF and science continuously dialogue and inspire each other.

2.3 Science fiction and science education

Given this relationship, the question of science fiction's potential role in science communication naturally arises. If SF is influenced by science, can it be used to teach or communicate science to the public? Science communication is multi-faceted and science communicators come from many backgrounds; not all forms of science communication are relevant to this thesis. The subset that is, however, is using science fiction for *science education*. Several studies analyse how SF affects its readership's understanding of science: Eichmeier et al. investigated audience understanding of the human genome (Eichmeier et al., 2023), Lowe et al. similarly investigated audience understanding of environmental science (Lowe et al., 2006), and Orthia interviewed *Doctor Who* viewers about how the show impacts their ideas about science (Orthia, 2019). There are examples of scientists using SF for teaching science to the public: the European Astrobiology Institute recommends its science fiction anthology *Life Beyond Us* for educational use (Nováková, 2023), several university courses use SF to teach science (Luokkala, 2014); (Saunders et al., 2004), and scientists have presented their research at SF conventions in 'science programming' tracks (Childers et al., 2023). Popular science books, which aim to teach general audiences about science, extensively cite fictional exoplanets to explain scientific concepts: Joshua Winn's *The Little Book of Exoplanets* references Tatooine to explain Kepler-16b, Jaime Green's *The Possibility of Life* references Solaris in talking about life in the universe, and Lawrence Krauss's *The Physics of Star Trek* uses the planet Vulcan to explain the history of planet detection in our solar system as well as using the *Star Trek* franchise more generally to introduce physics to readers. Analysing readership understanding and case studies of science fiction references are first steps; what is missing is a broader analysis of the scientific content of science fiction. This is the next step in the scholarship that my research

addresses.

Analysis of the scientific content of SF is needed because, although interest in its potential science education use has been established, SF often suffers from a reputation of being "lowbrow". A 2017 experiment found that readers of SF "exerted less inference effort for theory of mind, but more for understanding the world" and "performed more poorly on comprehension" (Gavaler and Johnson, 2017). Gavaler and Johnson's 2017 paper was widely reported and lambasted as the "sci-fi makes you stupid" study (Flood, 2019), and the authors later did another study, altering their methodology in response to critique, that found no difference in effort or comprehension between SF and realistic fiction versions of the same story. Since the story did not change other than one genre-defining word ("daughter" or "robot") they found that reading effort depended on the quality of the story, not the genre of the story. As in the 2017 study, they found that SF reading required more "theory-of-world" or worldbuilding effort (Gavaler and Johnson, 2019), in keeping with the genre's estrangements. That is, readers of SF exert effort to understand and consider the changes between the world of the story and the real world. The perception of SF as inherently lowbrow (that is, mass-market or intellectually undemanding) has likely contributed to its being overlooked as a source of science education, however, it is notable that the worldbuilding is the part of the story taking the most effort to understand from readers, because this is where the scientific inspiration can be found. I suggest that this focus on worldbuilding indicates SF may be a strong tool for teaching science in allowing readers to understand and consider the effects of science on the world. Over time, SF has moved from its relegated position with the pulp novels and B-movies, to achieve much more prominence in the 21st century, as evidenced by the audience statistics cited at the beginning of this chapter. For the purposes of my thesis, I am not interested in further addressing the literary quality of SF, but rather its scientific content. Subjective comments about "quality" have been a good motivator for a more rigorous scientific study of SF and of its scientific content, which my research addresses.

There is an established literature on the use of narrative and storytelling in educational settings, including in using stories to teach the history of science in science teacher education (Adúriz-Bravo, 2014), in secondary education (Mutonyi, 2015), and as a powerful pedagogy practice (Landrum et al., 2019). Mutonyi, in studying student use of stories in understanding the science of HIV in Uganda, found "stories, proverbs, and anecdotes can be used to help students who may not be inclined to science, to enter into the world of science by linking their

everyday world to the culture of science" (Mutonyi, 2015). Narratives are strong pedagogical tools because they create interest and provide structure for remembering information; they are also the oldest form of teaching known to humanity, dating back tens of thousands of years (Landrum et al., 2019). There is also evidence from neuroscience that reading stories aids in memory of information (Berns et al., 2013); Baldassano et al evidenced this by monitoring hippocampal activity while subjects watched and later retold the events of an episode of *Sherlock*, identifying the mechanisms of the strong memory production of storytelling (Baldassano et al., 2017). Irani looks specifically at fictional short stories as science education tools (Irani and Weitkamp, 2023), and Reinsborough's report of a 2017 panel at the UK Science in Public conference demonstrates interest in the use of SF in science education (Reinsborough, 2017). I extend this work to include the use of SF narratives, as these provide a "link...to the culture of science", as Mutonyi put it. Considering the accuracy of the SF story is important in considering its use in science education. Such uses would involve a science communicator to design lesson plans, training tools, or presentations around the SF story.

Several university courses use SF to teach scientific concepts to audiences largely composed of humanities students, building off of people's passion for the genre. SF can offer a common knowledge base: I frequently refer to a fictional planet (Tatooine) in my public engagement work to help audiences understand the term exoplanet, and anecdotally have found this to be common amongst astronomers. Barry B. Luokkala, who runs a science fiction and science education course at Carnegie Mellon University, makes the case that an understanding of science is necessary for citizens of our current technological world. He notes that most STEM courses for non-STEM majors in the United States fail to spark interest in learners because "students who take such courses often do so simply to fulfill a degree requirement and with a sense of trepidation. They often complete such courses with no more interest in science than when they signed up." His course uses SF's popularity and engagement factor to "make science accessible to a broad audience" (Luokkala, 2014). The University of Glamorgan in Wales developed an astronomy course taught using SF, and found they were able to engage populations it had been difficult to attract to higher education, like adult job-seekers without degrees (Saunders et al., 2004). They found through their modules that learners with no formal qualifications in science already had more scientific knowledge than expected, as well as a keen interest, from being fans of science fiction, using "knowledge sources that reach beyond the more conventional and obvious source literature of 'books'; extending into the critical reading of magazines, role-play

exercises, fanzines and even video games" (Saunders et al., 2004). Interestingly, this provides evidence that SF can achieve some science education without the formal intermediary of an educator. More recently, Stephen Kane's course at UC Riverside presents another example of teaching science through SF, and one specifically focused on exoplanets (Kane, 2023), and the University of New South Wales offers a course "Science in the Cinema" teaching general science requirements using science fiction (UNSW Handbook, 2024). These courses indicate that one advantage SF can provide for science education is that presenting science ideas through science fiction lowers barriers of academic hierarchies, privileges, and access. These courses have seen success in reaching wider audiences.

As the creators of such courses know, people like science fiction—and crucially, people who are not STEM majors enjoy reading and watching science fiction. Science fiction is no longer relegated to children's media or niche interest groups, but is today a powerful pop culture force full of symbols and stories and quotations recognised even by people who have never seen the film or read the book they originated from. As a review of a history-of-SF exhibit at London's Barbican—by its very existence an achievement in the mainstreaming of SF—put it, "sci-fi has become a logical means of making sense of an increasingly complex, interconnected world" (Lambie, 2017). In 2023, London's Science Museum also had a major exhibit called *Science Fiction: Journey to the Edge of Imagination*. By asking the reader to perform cognitive estrangements, science fiction puts our world in context. For reading and viewing audiences, seeing an exoplanet in science fiction can help them better understand their Earth by comparison. Much evidence has been presented suggesting SF's portrayals of science might have educative value—so, how does one study these portrayals? How does one scientifically analyse literature and media data for science content?

2.4   Scientific study of literature

There are precedents for using quantitative methods to study wider trends in genres, although traditional studies of literature tend to involve analysis of authors and characters and close reading—a technique focusing on a single story or story element. The Russian Formalists were the first to attempt the "scientific" study of literature. In the 1910s-20s, practitioners of the Formal Method argued for a more scientific study of literature that ignored the biographies and intentions of individual authors (Finer, 2010). They believed that poetic language, which they differentiated from the everyday, followed certain rules (Berlina, 2017). Indeed, Formalist

Viktor Shklovsky's concept of estrangement, or ostranenie, influenced Suvin's definition for SF (Berlina, 2017); (Slusser, 2003). Shklovsky posits that "the form of a work of art is determined by its relationship with other pre-existing forms…[Art is] created either as a parallel or an antithesis to some model" (Shklovsky, 1925). The development of literature is not linear, but rather works develop both from what came before and what comes alongside (Berlina, 2017); this connects back to the revolving door analogy and the circulating chain of reference. This type of thinking, combining the arts and sciences, was common in the Soviet Union at the time, as this 1924 *Pravda* quote suggests: "The creator of fantasy plays a visible role in science, as in art. Fantasy is absolutely necessary to the development of science, it is just as necessary as all the patient research collecting materials and observations. Without studying systematic facts, we have fruitless fantasising. Without fantasy, science turns into an accumulation of facts, classifications and inferences, it remains meagre and barren. It is in the harmonious convergence of scientific research and science fiction that we find a way to move forwards" (Lapirov-Skoblo, 1924).

The Formalists focused on form, and on literature as a dataset to be examined in volume. Thus, Vladimir Propp's analysis of a body of folktales allowed him to identify 31 basic structural elements present in the genre and show his findings through diagrams and graphical networks (Propp, 1968). More recently, Franco Moretti's work uses 'distant reading' and graphics to chart changes in genres over time, uncovering structures within books (Moretti, 2005). Current digital humanities work, such as Coll Ardunuy and Sporleder's work creating visual networks of character relationships in classic novels to show that such networks are different by genre, provide more examples of larger-scale studies of media (Coll Adanay and Sporleder, 2015). What these scholars share is that in losing the detail of individual works within their dataset they see the bigger picture, capturing wider trends in literature. Taking inspiration from these scholars, my research uses Bayesian network analysis, introduced in the next chapter, to study the portrayal of exoplanets in science fiction on a larger scale than has previously been done.

2.5 Conclusion

Science fiction is a genre that takes inspiration from science, often reflective of cutting-edge science at the time the work was written. SF writers and scientists have a mutual, reciprocal relationship, in which each group inspires the other, and often an SF writer and a scientist can

be one and the same person. This relationship stems from the role of empiricism in SF: authors have to make decisions about where to be realistic, plausible, or imaginative in their use of science. SF is popular, and often introduces its audience to scientific concepts—for example, many become familiar with the concept of exoplanets through science fiction stories—thus, it has potential uses for science education. Given that SF works are *fiction*, the challenge becomes how to delineate fact from fiction to use SF for science education and take advantage of the teaching power of narratives. For centuries, and in volume since the mid-20th-century, science fiction stories have featured exoplanets as settings. This prevalence of exoplanets in science fiction combines with recent discoveries in the field of exoplanet science to make SF featuring exoplanets an ideal dataset with which to investigate the portrayal of science in SF. Understanding the portrayal of exoplanets in science fiction, and how it has changed in light of scientific discovery, can help answer research questions about the science content of science fiction because a large dataset of SF exoplanets can be built. The next chapters will introduce the data science methodology I employ to answer these questions: Bayesian networks.

3

Approach and methodology

[Portions of this chapter are taken from my work in the paper 'Science Fiction Media Representations of Exoplanets: Portrayals of Changing Astronomical Discoveries' published in the March 2024 issue of JCOM Journal of Science Communication, and my paper 'Dialogues Between Science and Fiction in the Creation of an Anthology' published in BSFA Vector issue 297 (2023)]

My research is made up of one large and two smaller projects: Bayesian network analysis of fictional exoplanets, the St Andrews Centre for Exoplanet Science (StACES) anthology project, and the N=1 Problem interview project. This chapter presents the research methodologies used for each project. I introduce Bayesian network analysis as a tool to understand fictional exoplanet portrayal in science fiction, explain how fictional exoplanet data were collected to create my database, and go over the qualitative data collection methodologies for the StACES anthology and N=1 projects, both of which involved recruiting and interviewing participants.

All projects are linked through the overarching goal of investigating how humans envision exoplanets, especially in science fiction and dealing with the unknowns in the field. The

fictional exoplanets project compiles a database of science fiction representations of exoplanets, which is mined for trends in which planetary characteristics are interconnected in the literature, and how portrayal of fictional exoplanets evolves with evolution of scientific knowledge, such as the discovery of 51 Pegasi b. The StACES anthology project paired scientists and creative writers to jointly create science fiction works based on real research—an exercise in what science fiction (SF) might look like if a scientist is involved throughout the whole creative process. The data collected in the form of participant questionnaires throughout the writing allow for analysis of the decision-making process of how science is to be portrayed in the fiction. Finally, the N=1 project unpacks an ever-present element of all human work, both scientific research and fictional imaginings, relating to exoplanets: our low sample size of well-studied planets (only those in the solar system, and particularly Earth) with which to compare our new discoveries. Particularly, for those interested in studying or exploring the idea of life in the cosmos, the fact that we only know of one example of a planet with life is an important research consideration—especially as currently, the technology does not exist to detect an "Earth twin", an Earth-mass exoplanet orbiting at 1 AU (Earth-Sun distance) from a Sun-like star. Combined, the data from these projects reveal a less-studied side of exoplanet science: the links between the science and human imagination. SF becomes a data set analysed with data science as well as qualitative methodologies, and the processes of scientists and writers regarding how they make decisions about the portrayal of science in SF, and how they handle unknowns in their own fields, are recorded.

3.1 Database of fictional exoplanets

Research Questions: How can we study whether science fiction reflects scientific discoveries? Exoplanet science has made massive discoveries in recent decades. Are these findings reflected in science fiction exoplanets, and if so, how?

In this section, I briefly describe how a fictional exoplanet database was compiled (covered in-depth in Chapter 4), then in the next I provide a background on Bayesian networks, and explain the methodology used for learning Bayesian networks from the dataset.

The fictional exoplanets dataset is a collection of 212 fictional exoplanets from works of science fiction across different media. Planets were added to the database first from sources familiar to the author, usually popular SF franchises or classic books like *Star Wars*, *Star Trek*,

Dune, *Solaris*, etcetera. To avoid bias towards only works I had encountered, I also used a crowdsourced Google form that collected fictional planet data from anonymous submissions. The form encouraged people to submit as much or as little information as they could verify about the planet, asking after:

- name

- work of origin

- author

- medium

- publication date

- whether it is a moon or a planet

- mass

- radius

- planet type (e.g. gaseous, rocky, ocean, etc.)

- orbital distance

- number and type of stars it orbits

- whether it is a fictional portrayal of a real star system

- how many moons it has

- year length

- day length

- whether it has native life

- whether it has intelligent native life

- whether it has an atmosphere

- whether humans can breathe the atmosphere

- whether it is a single-biome planet

- whether it is a fictional portrayal of a real exoplanet

Lastly, there was an open-ended question to provide further information if desired. The form was shared in Twitter groups of astronomers and SF fans, and at conferences like the World Science Fiction Convention. Fictional exoplanets submitted were fact-checked and verified. In practice, most submissions did not contain answers in all fields, as very few fictional exoplanets have all these data available.

Each fictional exoplanet was added to our database and categorised according to nine variables, listed here but presented in full detail in Chapter 4, which describes the database. The variables were:

- AfterDiscovery (measured as before or after 1995)

- HabZone (measured as whether or not the planet is at the right distance from its star to have liquid water on the surface)

- RealStar (measured as whether the planet's star is real or fictional)

- Life (measured as whether the planet has native life)

- Intelligent (measured as whether the planet has intelligent native life)

- HumansBreathe (measured as whether humans can safely breathe the atmosphere)

- MediaType (measured as film, book, TV show, video game, or podcast)

- IsGas (measured as whether the planet is gaseous or terrestrial/rocky)

- EstNonNativeHumans (measured as whether the planet is home to an established, multi-generational human settlement)

The resulting database constituted the measured data that I input into my Bayesian network software. Although there are SF universes with hundreds of fictional worlds, biases were mitigated by not including large numbers of worlds from the same property. The highest number of planets from a single source is nine from the *Star Trek* franchise.

While attempts were made to include planets from a variety of time periods and types of media, I recognise a current bias in the data toward Western, English-language fiction. To best

answer research questions about the effects of new scientific discoveries on science portrayal in SF, an effort was made to collect reasonably equal numbers of fictional exoplanets from before and after 1995, the year in which the first real exoplanet was discovered orbiting around a Sun-like star. A full exploration of the fictional exoplanets database can be found in Chapter 4 and the full dataset can be accessed from the link in the Appendices.

3.2   Bayesian network

Research question: Bayesian networks are graphical representations of interconnected variables. What interactions between portrayals of fictional exoplanets and scientific knowledge can a network reveal?

Bayesian networks are graphical representations of a joint probability distribution—that is, the probability of multiple events happening together—of a given set of variables (Heckerman et al., 1995). They are presented as directed acyclic graphs (DAGs) where arrows between variables show their statistical dependencies (Pearl, 1988); (Smith, 2010). They are part of Bayesian statistics, founded on Bayes' Rule, which describes the conditional probability of an event given conditions relevant to the event. Mathematically, Bayes' Rule is:

$$P(A|B) = P(A)\frac{P(B|A)}{P(B)}, \tag{3.1}$$

A and B are generic events. Equation 3.1 can be read as the probability of A given B equals the probability of A times the probability of B given A, all divided by the probability of B. Bayesian networks evolved from decision-making systems that incorporated knowledge from human experts. Once an expert-informed Bayesian network was presented, Bayes' Rule was used to determine the values of some variables given the values of other variables. Bayesian networks are a versatile data science tool for trend mapping, and can aid in decision-making—this methodology has been applied to fields from genetics (Videla Rodriquez et al., 2022) to ecology (Hui et al., 2021) to astronomy (Pichara and Protopapas, 2013).

Structure learning of a Bayesian network—that is, learning the structure of the network that best fits the data, as described below—is ideal for a situation in which data have been measured, and one wants to find out how they are connected—exactly the situation with the fictional exoplanets database. In keeping with the trend-mapping approach of the Formalists and Moretti described in Chapter 2, I created a database of fictional exoplanets and employed

Bayesian network analysis to reveal the connections between the nine categories (variables) in that database, listed in the previous section. From this dataset, a search algorithm explored the space of possible networks and determined a solution that represents the patterns among the measured variables.
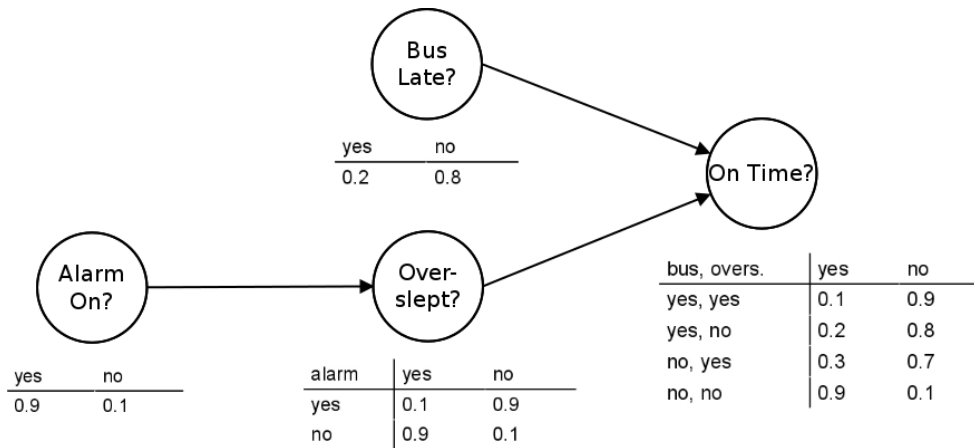
Figure 3.1: A simple example of a Bayesian network. The tables beneath each variable are conditional probability tables relating the values of the variable to the values of its parents. (Credit: Pekka Parviainen)

A Bayesian network is a graph (shapes with arrows between them) made up of *variables* and *links* between variables. Variables must be measurable; the property that is measured is what is influencing or not influencing other variables. To read graphs, a family analogy is used: if variable A influences variable B, A is B's *parent* and B is A's *child*. Figure 3.1 provides a simple example of a Bayesian network. Variables represent events like whether or not a person's alarm is on, whether or not the bus was late, and whether or not the person overslept. In the Bayesian network representing these variables, whether or not the person overslept is a parent of whether or not the person was on time. The network indicates which variables are useful for predicting others; the DAG shows which variables influence each other directly (if linked) and indirectly (if linked through several variables). The structure of the Bayesian network also reveals relationships such as *conditional independence*: a variable is conditionally independent of its non-descendants, given its parents. Conditional independence is the case where, given one variable, another variable provides no further useful information. For example, whether or not the alarm is on is conditionally independent of whether or not the person was on time, in Figure 3.1. Whether or not somebody is on time is dependent only on whether or not the bus is late or whether they overslept (which sits between the variables for Alarm On? and On Time? in the figure): whether they slept through the alarm or forgot to put the alarm on at

all does not matter. A joint probability distribution for the Bayesian network in Figure 3.1 can be written thusly:

$$P(A, B, C, D) = P(A)P(B|A)P(C)P(D|B, C), \tag{3.2}$$

Where A is Alarm On?, B is Overslept?, C is Bus Late?, and D is On Time?.

Bayesian networks allow analysis of SF as a complex interconnected system. Given a dataset where each data point is a fictional exoplanet and its associated values for a set of variables (representing characteristics of the planet, such as atmosphere type, and fully discussed in Chapter 4)—the fictional exoplanets database—I used a Bayesian network analysis software called Banjo to find a Bayesian network that best describes the statistical dependencies in the dataset, i.e. how each variable influences the others (Hartemink, 2010). My variables are discrete (they take on discrete value states). Bayesian networks, due to their probabilistic nature, can handle noise (in my case, uncertainty in exoplanet variables), and the use of discrete variables makes the network even more robust to noise. Because it is not possible to simply calculate from a dataset the Bayesian network that describes its joint probability distribution (Chickering, 1996), a form of unsupervised machine learning was used. This is *structure learning* of a Bayesian network. Banjo, the *search* algorithm, explored potential sets of links between variables and determined a network that best represents the statistical dependencies among measured variables in the data. Banjo *scores* networks with the Bayesian Dirichlet equivalent (BDe) score. The BDe score is one of a number of scoring metrics for Bayesian networks, which calculate the probability of a network given a presented dataset. The BDe score is for discrete data, and finds the probability of a network by making an estimate of the marginal likelihood integrated over the parameter space of all possible conditional probability tables. The BDe score provides an inherent penalty for overcomplexity, making it more suited for lower amounts of data (like my 212 planets) than other scoring methods like the Bayesian information criterion (BIC) (Heckerman et al., 1995), (Yu et al., 2004). (For a derivation of the BDe metric, see Heckerman.) Bayesian networks provide an advantage over running simple statistics like linear regression by enabling modelling of the entire context of measured variables and how they interrelate. In my case, subtleties of interactions among features of envisioned exoplanets can be placed within the larger context of the network and both direct (two variables linked in the network) and indirect (variables connected via a chain of

links) connections can be explored. These can then be used to infer about causal relationships between variables (Heckerman et al., 1995).

A *search* in Banjo starts with a random network (that is, set of links between the variables) and scores it against the measured dataset; it then makes a single change to the network (adding or deleting a link), and asks if the score of this new network is better or worse than the one before. By many such individual steps, millions of networks are explored in each search and the structure of the Bayesian network that best fits the data is learned (Heckerman et al., 1995). Bayesian networks come with links in the form of arrows (the 'directed' in DAG) indicating a statistical directionality that can be read as 'is useful for predicting'; there can be multiple networks that are *mathematically equivalent* where the usefulness for predicting can go in either direction (e.g., clouds can predict it might rain, but rain also predicts existence of clouds). With directionality, the variable the arrow is emerging from is the parent of the variable the arrow goes towards; however, networks that are mathematically equivalent can have arrows going in different directions, these are known as members of the same equivalence class. All Bayesian networks can be written as an equation of the joint probability distribution, as in equation 3.2. By algebraic manipulation substituting any term with its representation from Bayes Rule, one can transform any member of an equivalence class into another member of the same equivalence class. Banjo generates a network structure along with an *influence score* for each link, which measures the strength of the relationship between variables as well as whether it is positive or negative. Strength of influence scores is indicated by distance from zero. Influence scores can also be reported as non-monotonic, meaning there is no specific direction to the relationship, e.g., where a high value of one variable maps to both high and low values of another, and in this case the influence score is reported as 0.00 (Yu et al., 2004). The influence scores are influenced by the parents, and therefore different members of the same equivalence class can have different scores on the same link.

There are several search decisions to be made when using Banjo. Banjo can produce as a solution either a *top-scoring network* or an average of a user-defined number of top-scoring networks. Banjo supports two types of search algorithm, *greedy* and *simulated annealing*. Within these are two different ways of making individual steps between networks, and it further allows the user to define the period of time a search takes (longer searches explore more networks). Greedy search always picks the higher-scoring option when changing links to try a new network (Jungnickel, 1999); this means it always climbs to the nearest 'peak'—a network

where all networks one step (one single change) away score lower—from its random starting point. Simulated annealing search has a decreasing probability of choosing the lower-scoring option (Kirkpatrick et al., 1983). Simulated annealing can take a longer time to find high-scoring networks than greedy, but can also find high-scoring networks surrounded by lower peaks, whereas greedy would stop at the low peak and never go down and back up to find the higher one. The peak metaphor is visualised in Figure 3.2. When taking individual steps in the search, Banjo allows users to choose either *random local moves* (makes random changes in network links and takes the first better-scoring option) or *all local moves* (checks all possible changes and goes to the option which most improves the score). Random local moves can enable a broader search of structure, as it goes through more different networks faster, but all local moves ensures that a nearby higher-scoring structure is not missed. In my searches (see below), I varied the search algorithm and move type.

Bumpy Surface Plot
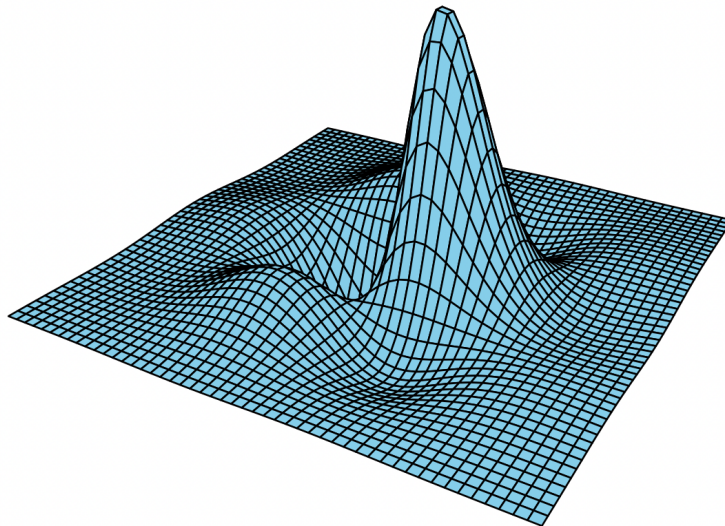
Figure 3.2: A representation of an example search space as a landscape with peaks. A greedy search always makes changes that increase the score (moving uphill in this metaphor). For a very bumpy solution space, greedy searches run a risk of getting "stranded" on a low peak. For a simpler solution space, greedy search is the fastest and most efficient search method to use.

The only other parameters that need to be set, and which I did not vary, are the *equivalent*

sample size and *best networks are*. The equivalent sample size reflects the level of weight to give to a Bayesian prior—a belief that any given link is present—in relative terms to data points (here, fictional exoplanets); a low value so that your data provides the bulk of knowledge is generally used: I picked 1, which becomes the imaginary sample size used by the BDe score and in this case indicates an even prior distribution. The even prior has a weight equal to one data point. The best networks can be set to either allow members of the same equivalnece class in a single search, or to ensure that all networks reported in a search represent different statistical descriptions of the data; the latter is better and only avoided for computing time reasons: I had no issues with computing time thus set the best networks to be nonequivalent.

Given all these trade-offs in the user-defined settings, I varied various parameters as described below to determine the best solution to use. To find the best settings, I ran 48 searches in Banjo. I explored (a) greedy and simulated annealing search, (b) random local moves and all local moves, and (c) allowed the search to run for 1, 5, 10, and 15 minutes, for a total of 2x2x4=16 different parameter combinations, and I repeated each combination for 3 searches. I used this to define the best search to use.

After running these searches, results were examined using BayesPiles, a visualisation tool that assists in determining a 'best answer' network by allowing researchers to explore solution spaces (Vogogias et al., 2018). BayesPiles analysis was used to determine the complexity of the solution space and therefore select the best and most efficient search methodology for our network. BayesPiles allows for direct comparison of networks found by each search, providing visual representations of how the links differ or are the same, and can be used to pile networks on top of each other in a matrix to view their similarities and differences.

Different searches resulted in top networks all of the same equivalence class. This was a clear signal that a top network and not model averaging is appropriate, and that no further, longer, searches need be run. The greedy, random-moves, 15 minute search was chosen to represent the final network because it is less computationally taxing than the simulated annealing search, which also turned up an equivalent result as revealed by BDe score.

Each search produced a single top network. Though these are all of the same equivalence class and this means they had links between the same variables, the reported directions of the links did change. For example, a link from variable A to variable B in one top network could become a link from variable B to variable A in another. The networks are mathematically

equivalent, as shown by all of my resulting top networks having the same Bayesian Dirichlet equivalent (BDe) score, a measure of how well the network fits the data. My final network includes the links without direction because of these differences among networks of the same equivalence class; further, I am not claiming causality but instead informativeness, and an undirected interaction better reflects this property. Once I chose my final search parameters (greedy, random-local-moves, 15 minute search length) I ran this search a total of five times to find multiple options of the same equivalence class provided by the top networks. This was because the influence scores reported in the resulting top networks can be different in different equivalence classes.

The final Bayesian network result was visualised with GraphViz (Gansner and North, 2000) to show the links between the variables for the top graph, as well as their influence scores, in order to examine the influence scores in the network.

The results of the Bayesian network analysis can be found in Chapter 5.

3.3 St Andrews Centre for Exoplanet Science Anthology

Research Questions: How are science fiction worldbuilding decisions made? What is the role of a science consultant in science fiction?

My production of a science fiction anthology entitled *Around Distant Suns: Nine Stories Inspired by Research from the St Andrews Centre for Exoplanet Science*, in which stories were created by teams of writers and scientists, presented an opportunity to gather data on how science decisions are made in the SF writing process. In creating SF stories, writers often will undertake their own background research, and sometimes will reach out to scientists to advise on the science content of the stories (see discussion on science and SF in Chapter 2, and background section in Chapter 6). The Anthology research project aimed to pair writers with dedicated science consultants who were to remain involved in the story creation process through to its completion. The research took some inspiration from art/science collaborations like St Andrews' *Intersections* project with the Royal Conservatoire of Scotland in which music is composed based on PhD theses, but envisioned a much more active role for the scientist, beyond merely acting as inspiration. The goal was to gain insight into the process of collaboration and decision-making between scientists and SF writers in the creation of SF works, and to help determine whether and in what role SF might contribute to science education.

Research Questions: What is the role of a science consultant for a SF story? How are decisions made regarding the portrayal of science in SF? What is the role of SF in science education?

Research ethics approval was obtained from the St Andrews University Teaching and Research Ethics Committee (UTREC) for the collecting of questionnaire responses from research participants. A copy of the approval letter for St Andrews UTREC number PA15192 can be found in the Appendices. It was possible to contribute to the Anthology without participating in the research; five of the nine pairs of participants in the Anthology participated in the research portion. Scientist participants were all scientists affiliated with the St Andrews Centre for Exoplanet Science. While the Centre is interdisciplinary and includes researchers in the fields of Modern Languages, Philosophy, and International Relations, among others, all the researchers who participated in this project were in natural sciences fields, including Physics and Astronomy, and Earth and Environmental Sciences, to best facilitate addressing the research questions. Scientist participants included faculty and postgraduate students. Writer participants were all postgraduate students, pursuing either Masters' or PhD degrees in English or creative writing at the University of St Andrews.

Scientists and creative writers were paired based on an interest form in which they indicated subject area preferences and what they hoped to gain from the project experience. Each team was instructed to meet at least three times over the two-month writing period in early 2021, and after each meeting each individual was to submit a filled-in questionnaire in which they detailed what had been discussed. The questionnaires included open-ended questions and prompts to facilitate broad responses. Questions were designed to focus on decision-making processes related to story development, in order to investigate the role of science in these processes. A link to the data in the form of pseudonymised questionnaires is available in the Appendices.

The Questionnaire asked:

What was discussed at this meeting?

What story ideas were generated?

What story decisions were made?

Describe any communication difficulties.

Describe any communication successes.

Describe the current status/progress of the story.

Teams were instructed to write a story based at least in part on the scientist's work, and that the scientist should be involved in the story creation beyond the initial story idea. Aside from those instructions, the details of the story-creation process were left to each team. This was in order to make the story creation process as natural as possible, and the research element as unobtrusive as possible.

All participants were full-time students or faculty during the course of the research, and the demands of schedules as well as the COVID-19 pandemic affected how often teams were able to meet. Not every team met three times, and not every participant submitted a questionnaire from each meeting. Results are based on six questionnaires from Team 1, three from Team 2, four from Team 3, four from Team 4, and seventeen from Team 5.

After the writing period, the surveys were collected and collated, with writers and scientists pseudonymised. Common trends and experiences were identified, as well as conclusions regarding the role of SF in science education and limitations of the study. These are detailed in Chapter 6.

3.4   N=1 Problem Project

Research Question: In what ways does the N=1 Problem affect exoplanet science research?

The N=1 project is a study of 18 researchers in the field of exoplanet science asked about the significance of the "N=1 Problem", that is, the fact that we only know of one planet with life (Earth), on their research. Exoplanet science research focuses on finding and characterising the diversity of planets that exist. While finding life is not explicitly a goal of exoplanet science, finding terrestrial planets in the habitable zone is often a goal of future missions, like PLATO (PLAnetary Transits and Oscillation of stars) (Rauer et al., 2014), and collaborative work on atmospheric biosignatures (potential indicators of life) with astrobiologists involves exoplanet scientists in the study of life in the universe (Fujii et al., 2018). Therefore, this study seeks to investigate exoplanet scientists' familiarity with the N=1 Problem and whether and to what extent it affects their exoplanet science work. This study is further motivated by the fact that, due to the N=1 Problem, scientific studies of life in the universe, including Search for Extraterrestrial Intelligence (SETI) research and astrobiology research, and imaginative portrayals of life on fictional exoplanets alike are affected by this "sample size problem" of having a single

example of life. While there exists a literature on the N=1 Problem, I am not aware of any prior use of a qualitative social science interview methodology to investigate how it pertains to exoplanet science research.

This study lies at the junction of several research areas: interviews discuss astrophysics, astrobiology, and earth sciences, while the interview methodology contributes to science and technology studies, ethnographies of scientists, and futures studies. For a more detailed background on this motivation, see Chapter 7. The goal of the project is to begin investigating the range of thoughts, opinions, and areas of opportunity, concern, or challenge relating to the N=1 Problem as seen by current practitioners in the field of exoplanet science. Due to small sample size of interviewed scientists, this is a preliminary study that is only generalisable as ways exoplanet scientists *can* respond to the N=1 Problem, rather than all the ways that they *do*.

Research ethics approval was obtained from the St Andrews University Teaching and Research Ethics Committee (UTREC) to interview participants who are researchers in the field of exoplanet science. The approval letter for St Andrews UTREC number ML16820 can be found in the Appendices. A Call for Participants was circulated through institutional mailing lists and the Exoplanet UK Community Mailing List (https://lists.cam.ac.uk/sympa/info/physics-exoplanets-exouk).

Researchers at various stages of their careers were interviewed, in order to capture a variety of experiences within the still-new field of exoplanet science. 18 participants were interviewed in spring 2023. Participants were current researchers at institutions in the United Kingdom, Austria, and the United States, but represented a wider range of nationalities, with many having lived and worked internationally. 8 were women and 10 were men. 6 of the 18 interviewees were current or former members of the St Andrews Centre for Exoplanet Science. Participants represented a range of academic experience, from PhD students to professors. All but one (a geoscientist) were astronomers, working on a range of exoplanet science topics including protoplanetary disks, astrochemistry, atmospheres, bulk composition characterisation, binary star systems with exoplanets, hot Jupiters, planet formation, population studies, Earth as an exoplanet analogue, directly imaged planets, and ESA and NASA mission design. Theory and observation were both well-represented in this range of topics.

Interviews lasted between 15 and 35 minutes and were conducted over Microsoft Teams,

with transcripts recorded (a link to the transcript data is available in the Appendices). Participants were not given access to the questions beforehand in order to facilitate extemporaneous responses. Before the interview, participants were informed that they could spend as much or as little time answering a question as they needed, and that "I don't know" was an acceptable answer. All participants were asked the same questions in the same order, and were given the opportunity to elaborate on any of their answers at the end of the interview.

Questions were selected to address the participant's research, their familiarity with the N=1 Problem, their perceived relevance of the N=1 Problem to their research and the wider field of exoplanet science, and their emotional response to the N=1 Problem. Questions were open-ended by design to facilitate discussion by participants.

The questions asked were as follows:

What job title do you identify with?

What is your work?

How does your work relate to exoplanets?

How does your work relate to the search for life in the universe?

What is your familiarity with the N=1 problem? Do you call it by another name?

How do you encounter the N=1 problem in your work?

Do you consider the N=1 problem to be a problem?

Does the N=1 problem hinder or help your research, or both? How?

What would it mean for your work if life were found elsewhere in the solar system? What if it were found on an exoplanet?

Do you believe that, with current technology, we can unambiguously detect life on another planet?

How do you manage the N=1 problem in your research? Are you happy with how you handle it?

Do you think astrobiologists and exoplanet scientists overall manage N=1 well? How could they improve?

How does it make you feel? Does it bring up any emotions?

Participants were pseudonymised and the interview transcripts analysed for common trends and experiences. Results indicated a wide range of both perceived relevance of and responses to the N=1 Problem among the 18 researchers interviewed. Results are detailed in Chapter 7.

3.5   Conclusion

This chapter has introduced the methodologies of the three research projects that make up this thesis. Methodologies have been presented and data used are provided in the Appendices such that all research could be replicated. Research uses a combination of quantitative and qualitative procedures, and includes an interdisciplinary form of knowledge production through the use of quantitative data science in the form of Bayesian network analysis of a media dataset. This mixed-methods research expands previous case study research into the portrayal of science in science fiction to investigate a larger dataset. While this is in keeping with Moretti's 'distant reading', the StACES Anthology study includes the thought processes of individual writers. The N=1 Problem study investigates an important consideration in exoplanet research and science fiction portrayals of exoplanets alike. These projects combine to investigate the links between scientific knowledge, the unknown in science, and science fiction.

4

Creating a database of exoplanets in science fiction

"Then what are You, having no Chaos found

To make a World, or any such least ground?

But your Creating Fancy, thought it fit

To make your World of Nothing, but pure Wit.

Your Blazing-World, beyond the Stars mounts higher,

Enlightens all with a Cœlestial Fier."

      —William Newcastle, *To The Duchesse of Newcastle, On Her New Blazing-World*

[Portions of this chapter are taken from my work in the paper 'Science Fiction Media Representations of Exoplanets: Portrayals of Changing Astronomical Discoveries' published in the March 2024 issue of JCOM Journal of Science Communication]

Exoplanets have featured in science fiction long before they were detected by astronomers—and before the word "exoplanet" was widely-used in the English language (Michel et al., 2011) (see

Figure 4.1). In science fiction (SF), they are often simply called "planets"—albeit ones outside our solar system, thus making them exoplanets. I have constructed a database of 212 exoplanets from works of science fiction. The goal was to analyse this database as a Bayesian network, which is done in Chapter 5, but the database is also an interesting result in and of itself. While there have been some academic investigations of SF exoplanets, and attempts to compile SF locations, previous academic efforts have not been very extensive or quantitative. In 2014, Stephen Baxter wrote a paper on Super-Earths—rocky planets more massive than Earth—in science fiction. While his paper represents a useful analysis of SF exoplanets that are of a different class from our solar system examples, he only investigates about a dozen worlds, and does not employ data science methodologies (Baxter, 2014). There also exist popular works like Stableford's *The Dictionary of Science Fiction Places*, which contain large numbers of fictional exoplanets, but they are targeted towards SF fan audiences and have widely-varying qualitative descriptions only: for example, some entries contain only a paragraph on the political situation on an exoplanet; others contain full numerical specifications (Stableford, 1999). My fictional exoplanets database represents an effort to collect a larger sample of exoplanets, and categorise them such that the same set of data is available for each entry.

I present this database in this chapter, detailing my choice of categorical variables, and providing examples of how I categorised planets. I also present overall statistics of my fictional exoplanets database, which, due to the often human-centred narratives of SF stories, is overall much more Earth-like than the population of real exoplanets, as will be shown.
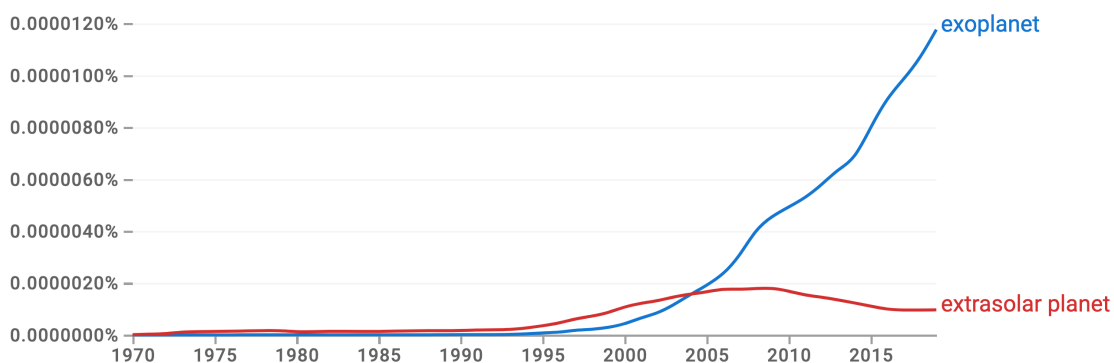
Figure 4.1: Google N-Grams chart of usage frequency of "exoplanet" and its currently-less-favoured alternative "extrasolar planet" in a corpus of books published in English in all countries, 1970-2019. Both terms hardly saw any use before the 1995 discovery of 51 Pegasi b.

Table 4.1: The variables for the fictional exoplanets database. The liquid water habitable zone refers to Kopparapu's (see text) definition "traditionally defined as the circumstellar region in which a terrestrial-mass planet with a CO_2–H_2O–N_2 atmosphere can sustain liquid water on its surface".

Variables for Fictional Exoplanet Database

Variable Name	Description
AfterDiscovery	Whether the planet first appeared in fiction before (0) or after (1) the discovery of real-life exoplanets around sun-like stars in 1995.
HabZone	Whether the planet is shown to be in the liquid-water habitable zone (1) or outside it (0).
RealStar	Whether the planet is portrayed as being part of a real star system (1) or not (0).
Life	Whether the planet is home to native life (1) or not (0).
Intelligent	Whether the planet is home to intelligent native life (1) or not (0).
HumansBreathe	Whether human characters can breathe the planet's atmosphere without assistance or ill effect (1) or not (0).
MediaType	Whether the planet originally appeared in a film (0), a book (1), a TV show (2), a video game (3), or a podcast (4).
IsGas	Whether the planet is Earth-like and rocky/terrestrial (0) or Jupiter-like and gaseous (1).
EstNonNativeHumans	Whether the planet has been colonised by an established population of non-native humans who have been there for hundreds or thousands of years (1) or not (0).

4.1 Variables

When planets were entered into the database (after being submitted to the Google form and validated for accuracy, as described in Chapter 4), they were categorised according to nine variables. These variables—AfterDiscovery, HabZone, RealStar, Life, Intelligent, HumansBreathe, MediaType, IsGas, and EstNonNativeHumans—are described in Table 4.1.

Variable creation was designed to include binary characteristics (binary variables aid in visual comprehension of the connections in the resulting network) that address scientifically-relevant statistics about the fictional exoplanets and the process by which they were developed. For example, planet type is represented with the variable IsGas. While proposals for exoplanet classification schemes exist (Kopparapu et al., 2018), there is not currently a 'gold standard' and planet types such as terrestrial planets, Super-Earths, Sub-Neptunes, Hot Jupiters, and

more, do not currently have precise definitions in science—unlike, say, the well-established stellar classification system which is based on spectral characteristics. Still, IsGas is a useful binary simplification of previously-known solar system planet types and newly-discovered exoplanet types into whether a planet's surface is rocky or gaseous, grouping fictional Super-Earths with the rocky planets and fictional Sub-Neptunes with the gaseous ones.

Because a lot of SF deals with the possibility of life in the universe (and because this has a relationship with exoplanet science searches for terrestrial planets in the habitable zone, see Chapter 7), many variables deal with fictional life on exoplanets: whether a planet has life and if it is intelligent, whether a planet's atmosphere can be breathed by humans, and whether the planet is in its star's habitable zone. While the concept of "habitable zones" did not enter wider use until after the discovery of exoplanets, I have included it because the concept does long predate the discovery of exoplanets (Lingam, 2021), appearing in Maunder's *Are the Planets Inhabited?* (Maunder, 1913), and is easily retroactively applied to older SF works. The habitable zone can have different meanings across the astronomical literature; here I use "traditionally defined as the circumstellar region in which a terrestrial-mass planet with a CO2–H2O–N2 atmosphere can sustain liquid water on its surface", that is, the distance from a star at which a planet can have liquid water on its surface given an Earth-like atmosphere (Kopparapu et al., 2013). It is this definition I apply to the fictional exoplanets in the database, with the caveat that an "Earth-like atmosphere" refers to atmospheric pressure rather than composition, so a planet being in the habitable zone does not mean human characters must be able to breathe. As well, fictional gas giants can be in the habitable zone, despite not having surfaces—some science fictional gas giants have habitable moons.

The variable RealStar communicates scientifically-relevant information because, for example, if a fictional planet is placed in a real star system, this shows the author must have done some research about star systems. Similarly, if the story were set on a real exoplanet, that would also evidence some level of research by the author. Some exoplanets in the database, such as Kepler-22b from the 2020 television series *Raised by Wolves*, are fictional portrayals of real exoplanets. However, I did not include a variable for whether the planet is a fictional portrayal of a real planet because this would necessarily only occur after discovery of real exoplanets, and thus be a subset of the intersection of RealStar=1 and AfterDiscovery=1. I did not feel this finer division of the intersection of these variables would add further information to our analysis, compared with RealStar. In contrast, Intelligent=1, despite being a subset of

Life=1, provides additional details, showing (particularly with Life=1, Intelligent=0) understanding of the potential for extraterrestrial life.

AfterDiscovery groups the fictional exoplanets into those that were worldbuilt before the author could have had any knowledge of real exoplanet discoveries, and those that were worldbuilt after such discoveries, and could have taken inspiration from exoplanet science findings. An effort was made to collect reasonably equal number of examples of exoplanets before and after the discovery of real exoplanets, to investigate the influence of scientific discovery on fictional exoplanet worldbuilding from the research questions.

MediaType is the only variable that is not binary, and allows for investigation of whether the medium of the fiction (film, book, television show, video game, or podcast) affects the other variables.

EstNonNativeHumans was a later addition to the list of variables for the fictional exoplanets database. When determining whether life on a fictional exoplanet was native or non-native, many worlds were encountered that were home to intelligent life only in the form of non-native humans. This occurred often enough to merit the addition of its own category. This is especially common in large SF universes in which humans have colonised a galaxy, like *Star Wars* or *Foundation* or Le Guin's Hainish Cycle.

Figure 4.2: The fictional circumbinary exoplanet Tatooine, as depicted in Star Wars Episode IV: A New Hope. Its two suns are visible.

4.2 Example fictional exoplanet classifications

As an example of this classification system, I demonstrate how variables were assigned for the fictional exoplanet Tatooine, shown in Figure 4.2. Tatooine is from the *Star Wars* franchise. It is a desert planet with a hot climate, located in the Outer Rim of the galaxy in which *Star Wars* is set. Notably, Tatooine is in a binary star system. The planet orbits two suns, one red and one yellow, of different stellar types (Lucas, 1977). Seeing Tatooine in *Star Wars* films was likely the first introduction of many members of the public to the concept of planets that orbit multiple stars, and perhaps to the concept of binary and multiple-star systems. In 2011, when the Kepler Space Telescope discovered a real circumbinary exoplanet in Kepler-16b, astronomers and science reporters alike frequently referenced Tatooine, the fictional appearance of which predated the first 1995 exoplanet discovery (Smithsonian Science, 2011). Anecdotally, in my science communication work, I have observed Tatooine to be the most frequently referenced fictional exoplanet when communicating science to the public.

Here is how Tatooine was entered into the database:

1 – AfterDiscovery: 0. Tatooine first appeared in Star Wars Episode IV: A New Hope in 1977, before the discovery of real exoplanets, and therefore is assigned 0 for AfterDiscovery.

2 – HabZone: 1. Although Tatooine has a hot and arid desert climate, its temperature range is temperate enough that it could support liquid water on its surface, therefore HabZone is 1.

3 – RealStar: 0. Star Wars takes place 'in a galaxy far, far away', so RealStar is 0.

4 – Life: 1. Several forms of life native to Tatooine are depicted, such as sarlaccs and dewbacks, so Tatooine gets a 1 for Life.

5 – Intelligent: 1. In addition to these non-intelligent species, Tatooine is home to native intelligent species like Jawas and Tusken Raiders, so it also gets a 1 for Intelligent.

6 – HumansBreathe: 1. Human characters such as Luke Skywalker are shown to breathe the planet's atmosphere without assistance, so HumansBreathe is 1.

7 – MediaType: 0. Tatooine originally appeared in a film, so despite its subsequent appearances across other forms of media, it gets a 0 for MediaType, which is defined by the first appearance of the planet.

8 – IsGas: 0. Tatooine is a rocky, terrestrial world, so IsGas is 0.

9 – EstNonNativeHumans: 1. There is an established population of non-native humans on Tatooine (in the Star Wars universe, humans are native to the planet Coruscant), so EstNon-NativeHumans is 1.

Figure 4.3: Melancholia about to collide with Earth. Image credit Melancholia (2011) by Lars von Trier.

As another example, I categorise the fictional exoplanet Melancholia, shown in Figure 4.3. Melancholia is a free-floating (not gravitationally tied to a star) gas giant planet in the 2011 Lars von Trier film of the same name. In the film, astronomers discover that Melancholia is on a collision path with the Earth (von Trier, 2011).

Here is how Melancholia was entered into the database:

1 – AfterDiscovery: 1. Melancholia is from a 2011 film.

2 – HabZone: 0. Melancholia is a free-floating planet, not orbiting any star and therefore not in the habitable zone.

3 – RealStar: 0. Melancholia does not orbit a star.

4 – Life: 0. Melancholia has no native life.

5 – Intelligent: 0. Melancholia has no native intelligent life.

6 – HumansBreathe: 0. Melancholia has an atmosphere similar to solar system gas giants, which humans cannot breathe.

7 – MediaType: 0. Melancholia appeared in a film of the same name.

8 – IsGas: 1. Melancholia is a gas giant planet.

9 – EstNonNativeHumans: 0. No humans live on Melancholia.
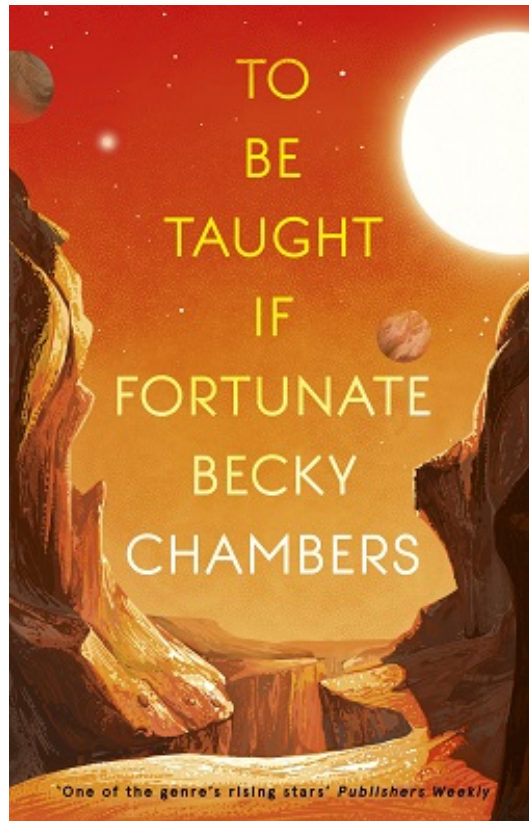
Figure 4.4: The cover of *To Be Taught, If Fortunate*

Finally, I present Mirabilis, from Becky Chambers' 2019 novella *To Be Taught, If Fortunate*, shown in Figure 4.4. Mirabilis is a rocky Super-Earth orbiting a red dwarf star. Human explorers travel to the planet and alter their bodies with a fictional process called somaforming that increases their muscle mass and bone density to better withstand Mirabilis's heavy gravity. Human explorers find a thriving ecosystem on Mirabilis, and speculate on whether it might actually be a more suitable host for life than Earth, perhaps a reference by the author to the concept of a "superhabitable planet" (Heller and Armstrong, 2014).

Here is how Mirabilis was entered into the database:

1 – AfterDiscovery: 1. Mirabilis appeared in the 2019 novella *To Be Taught, If Fortunate*.

2 – HabZone: 1. Mirabilis orbits its red dwarf sun in the liquid-water habitable zone.

3 – RealStar: 0. The red dwarf star this system orbits is fictional.

4 – Life: 1. The human explorers encounter a variety of plant and animal life.

5 – Intelligent: 0. The life on Mirabilis is not intelligent, though characters speculate one day it may become so.

6 – HumansBreathe: 0. Human characters use spacesuits to breathe.

7 – MediaType: 1. Mirabilis is from a book.

8 – IsGas: 0. Mirabilis is a terrestrial, rocky world, specifically a Super-Earth.

9 – EstNonNativeHumans: 0. The human characters in the novella visit Mirabilis, but there is no permanent human settlement there nor do they plan to create one.

4.3   Limitations and Decisions

The database contains a variety of fictional exoplanets, from both "hard" (more scientifically-accurate) and "soft) (less scientifically-accurate) SF sources, from the 1600s to 2023, and from a variety of forms of media. However, I recognise a bias towards Western, English-language fiction. Some planets from works in translation are present in the database, such as Trisolaris from *The Three-Body Problem* and Solaris from *Solaris*, but English language works remain overrepresented.

Data sources, and variable selection, were limited by the need to fill every category for every fictional exoplanet. While theory exists for handling missing data in Bayesian network structure learning (Friedman, 1998), there is a lack of software which provides sufficient implementation and interoperability of this with other tools (like BayesPiles, see Chapter 5) for gaining confidence in resulting network structures. Simple imputation methods, such as replacing empty values with the modal value, would skew the data too much. Therefore, I only included a planet in the database if I could answer every variable, thus excluding worlds only briefly visited in their SF stories.

In those cases where a fictional exoplanet's status has changed over time—for example, a planet which used to host life, but it has since died out—the planet's state at the time in which the main story is set is used. I had anticipated there might be some cases in which it might be difficult to determine whether a planet's life was intelligent or not; this ended up not being the case. A common plot device in SF is to feature human characters encountering alien life they do not initially realise is intelligent, but the reveal and confirmation of intelligence are

often included in the story. The *Star Trek* episode "The Devil in the Dark" is a good example of this. Janus VI, the fictional planet on which this story takes place and home of the intelligent Horta, is included in the database. Even in Lem's *Solaris*, in which the alien ocean is meant to be unknowable to the human characters, it is still clear that it represents an intelligence. Discrete judgement calls were made when needed regarding a reasonable level of ambiguity, and if a variable status was too ambiguous, the fictional exoplanet was not included in the database.

4.4   Database statistics and analysis

Figure 4.5 displays an overall description of the 212 fictional exoplanets included in the database. There is a near-even split of fictional exoplanets from before and after the discovery of real exoplanets, showing the attempt to collect equal amounts of examples for each was successful. Many of the planets have Earth-like features—more than half of the fictional exoplanets in the database are located in the habitable zone, are places where human characters can breathe, and are home to native life of some sort—with half home to intelligent native life. In fact, 59% of the planets in the database are terrestrial worlds in the habitable zone with human-breathable air. About 30% orbit real stars, and the same number are home to established non-native human populations. Only about 11% of the fictional exoplanets in the database are gas giants. While more than half of the fictional exoplanets come from literature, there is significant representation from video games, film, television, and podcasts.

The real exoplanet population is very different from the fictional population in this dataset, as demonstrated by Figure 4.6. Figure 4.6 charts the confirmed population of exoplanets by planetary mass versus orbital period, and I have added a green circle to show the position of Earth. To date, no exoplanet with the same mass and orbital period as Earth has been discovered, and the area of the chart both less massive than Earth and with a longer orbital period (the region above and to the left of Earth) is entirely devoid of data. Since specific masses and years of fictional exoplanets are rarely provided by authors, I cannot overplot my own fictional exoplanets database on Figure 4.6; however, I have added a few representative fictional exoplanets with such numbers provided, and comparing to Figure 4.5 as well shows that these populations are very different—Figure 4.5 has a majority of Earth-like planets, and Figure 4.6's population is mostly of much more massive planets orbiting closer-in than the habitable zone. Any exoplanet in Figure 4.6 more massive than 10 Earth masses (1e+1 on the
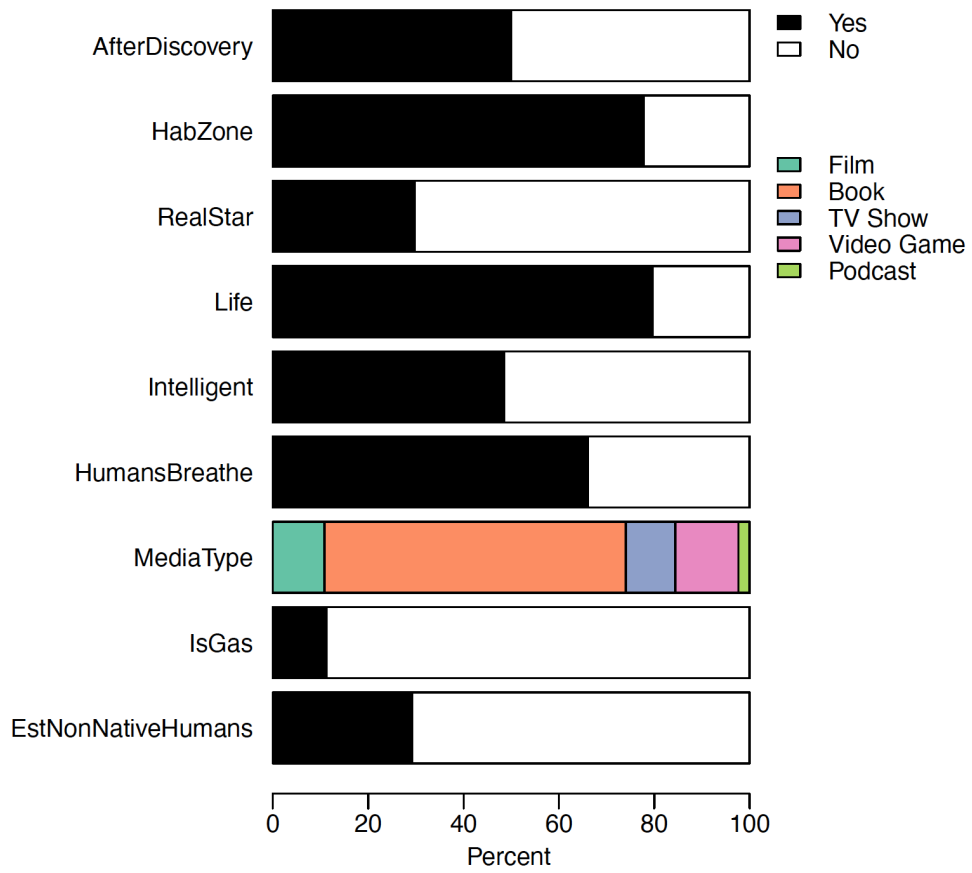
Figure 4.5: Statistics for each variable in the database.

logarithmic scale) is likely to be a gas giant, a much larger percentage than 11% of my fictional exoplanet database. While some of the exoplanets in Figure 4.6 *are* in their stars' habitable zones, these are planets orbiting stars much less massive than our sun that have closer-in habitable zones, like red dwarfs. Only a handful of fictional exoplanets in my database, like Superman's homeworld of Krypton and the four planets in Becky Chambers' *To Be Taught, If Fortunate*, orbit red dwarfs.

Given the youth of the field of exoplanet science, we are only just beginning to be able to perform population statistics on discovered planets and extrapolate about the overall exoplanet population, of which the 5000-odd planets discovered to date still constitute a tiny fraction. Commonly-used detection methods like the transit method are biased towards short-period, massive planets, and the instrument sensitivity to detect an Earth-twin does not yet exist. Though such planets are yet to be observed, there are studies that estimate occurrence rates of rocky planets in the habitable zones of Sun-like stars for purposes of mission-planning for future exoplanet characterising missions—Hsu et al. estimate a range of 0.03–0.40 such planets per star (Hsu et al., 2019). It is clear that most currently-discovered exoplanets are not human-habitable. (As a note, the upcoming Nancy Grace Roman Space Telescope will include a microlensing survey that will change known exoplanet demographics by detecting more exoplanets at large orbital separation (Johnson et al., 2020)). Of the 4700 planets in the Kepler confirmed and candidate population, only 7% orbit in the habitable zone, as defined by placing limits on solar insolation received and equilibrium temperature (NASA Exoplanet Archive, 2024), and being in the habitable zone says nothing about, for example, whether the planet has a breathable atmosphere. The fictional exoplanets database, in contrast, has more than half its worlds in the habitable zone. These comparisons demonstrate that the dataset of real exoplanets contains far fewer Earth-like worlds as a percentage than my fictional dataset.

There are several reasons for this. The obvious is that SF stories are written by humans, for humans, and often feature human main characters. It is narratively easiest to have these human characters exploring worlds where they are comfortable, and indeed, within the narrative, the human characters may purposely choose to visit more human-friendly planets. Another reason is medium-dependent—while a novel can easily depict any sort of world through prose, a live-action movie or TV show is dependent on budget and might portray Earth-like worlds due to this fiscal and practical limitation. This is why alien planets in *Star Trek* often look suspiciously like the California desert. As well, science fiction is, ultimately, fiction, and is not
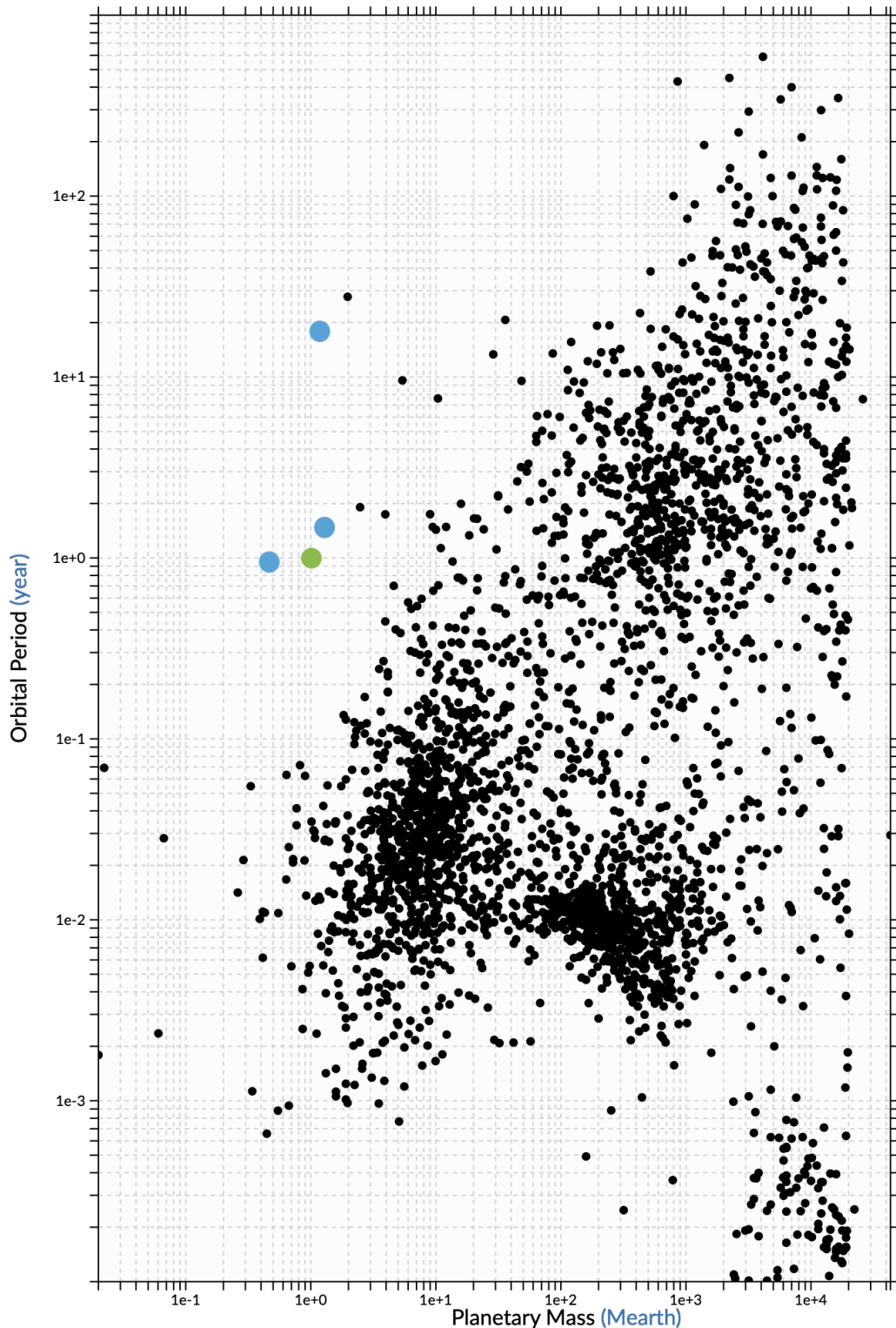
Figure 4.6: A chart showing a representative population of 2601 confirmed exoplanets. Planetary mass is plotted on the x-axis in units of Earth mass (Mearth). Orbital period is plotted on the y-axis in units of years. Axes are logarithmic. The position that Earth would occupy on the chart is indicated by the green circle, showing no exoplanet with the same mass and orbital period as Earth has yet been discovered. Several fictional exoplanets from my database are overplotted in blue: Tatooine (assuming it has Earth density), Klencory from *Mass Effect*, and Iris from *Aurora*. (Credit exoplanets.eu, modified with inclusion of green circle for Earth and blue for fictional exoplanets, licensed under CC BY 4.0 https://creativecommons.org/licenses/by/4.0/).

required to be scientifically accurate—there are physically implausible and impossible planets in the fictional exoplanets database, such as Goldblatt's World in *The Integral Trees*. The N=1 Problem, the fact that we know of only one planet with life on it and that is Earth, also likely plays a role in the prevalence of Earth-like exoplanets in science fiction. See Chapter 7 for more on this. At first glance, it seems therefore that the answer to the research question *Exoplanet science has made massive discoveries in recent decades. Are these findings reflected in science fiction exoplanets, and if so, how?* is no. However, for the reasons listed above, one cannot expect the population of fictional exoplanets to look anything like the currently-discovered population of real exoplanets (which is biased by detection methods), though comparing and contrasting them can still be a valuable exercise. I include this discussion of real exoplanets to put the changes in the fictional exoplanets over time in context. In order to determine if exoplanet science findings are reflected in science fiction exoplanets, one must compare science fiction exoplanets not to real exoplanets, but to themselves. This is done in chapter 5, in which the fictional exoplanets Bayesian network is presented and analysed.

Science fiction is a literature of ideas, of "what-ifs"—as discussed in Chapter 2, stories take our familiar world and change something, then see what happens. This places an emphasis on setting in the genre, and as a result many science fiction exoplanets can fulfill the functions of characters themselves in their stories. Sometimes this is literal: Pandora in *Avatar* has a planetwide consciousness called Eywa, Arrakis in *Dune* and Solaris from *Solaris* similarly exhibit planetwide consciousnesses. Clarke notes that *Dune, Neuromancer,* and the famous Earthrise photo from Apollo 8 showing the Earth viewed from lunar orbit are all part of a new planetary imaginary that "renovates intuitions of the actual Earth in its complex operations or that inspires new fictive worlds to refract those processes for us" (Clarke, 2015). In this way, fictional exoplanets represent ways for humanity to learn about ourselves and understand our own planet, Earth. Every fictional exoplanet is implicitly created to be compared to the Earth, and the familiar, by the reader. Stephen Kane, in his exoplanet course for non-science majors at UC Riverside, also emphasises this idea of "planet-as-character", explaining he uses this concept from literature to "meet [his] students halfway" in the course material (Kane, 2023). "Planet-as-character" and the use of fictional planets to hold a mirror to humanity and Earth may also help to explain the prevalance of more human-friendly fictional exoplanets.

The fictional exoplanets database is a significant corpus of SF planets, showcasing both the creativity of SF authors, and their tendency to worldbuild planets with Earth-like characteris-

tics. A few possible reasons for this have been presented here. The database is categorised and fully analysable by data science methods, as is done in Chapter 5. I want to conclude by referencing again Baxter's paper on Super-Earths in science fiction, since this is another attempt to understand how earlier SF portrayed a class of planet that was not yet known to exist before exoplanets were discovered. Baxter provides valuable examples of how SF can use the science of its day to construct fictional planets, and then be proven "wrong" by later observations and new knowledge. Many of Baxter's authors had, for example, calculated surface gravities for their fictional worlds. None, however, anticipated the true depths of oceans on waterworlds, and few of his author's planetary compositions and densities hold up. Yet, referencing the creativity of Lem's Solaris and Niven's Jinx (both in my database), he concludes "but the legacy of this sub-genre is perhaps more impressionistic. When trying to get the numbers right in the context of the science of the time, writers have intuited the unfamiliarity of these strange worlds, huge, varied, perhaps with surface conditions unlike anything we have previously experienced or imagined" (Baxter, 2014). The fictional exoplanets database is a repository of human understandings of other planets.

Bayesian network analysis: fictional exoplanets respond to scientific discovery

> *"My reason was convinced, answered the Marchioness; but my imagination is overwhelmed with such infinite variety and number of inhabitants existing in each of the planets; for as there is no dull uniformity in nature, the difference of species must be in proportion to the number of beings—how can imagination grasp such a vast idea?"*
>
> —Bernard Le Bovier de Fontanelle, *Conversations on the Plurality of Worlds*

[Portions of this chapter are taken from my work in the paper 'Science Fiction Media Representations of Exoplanets: Portrayals of Changing Astronomical Discoveries' published in the March 2024 issue of JCOM Journal of Science Communication]

This chapter presents the results and analysis of the Bayesian network created from the fictional exoplanets database. Significant links between variables are highlighted, and results are contextualised by discussing changes in the genre of SF and the known population of

real exoplanets. Overall, this Bayesian network analysis provides evidence that science fiction exoplanets are trending less Earth-like in response to scientific discovery of real exoplanets.

5.1 BayesPiles

BayesPiles (Vogogias et al., 2018) was used as a visualisation tool to explore solution spaces of the top networks of different searches. As a reminder, five searches were run using the selected parameters in Chapter 3 (greedy search, random local moves, 15 minute search length), in order to explore different options of networks of the same equivalence class, because influence scores are related to the parents and can be different across the same equivalence class. Influence scores are not represented in BayesPiles, but links and reported directionality are. Each search returned a top ten non-equivalent networks, and across the five searches the top networks (that is, each search's highest-scoring network) were all in the same equivalence class. Figure 5.1 compares the top ten non-equivalent networks found by each of the five searches. Although the top networks are of the same equivalence class, it is clear from the variations in shading (out-degree) that the paths each search took to arrive at a top network were different, and that the top networks themselves have some slight variations in directionality.
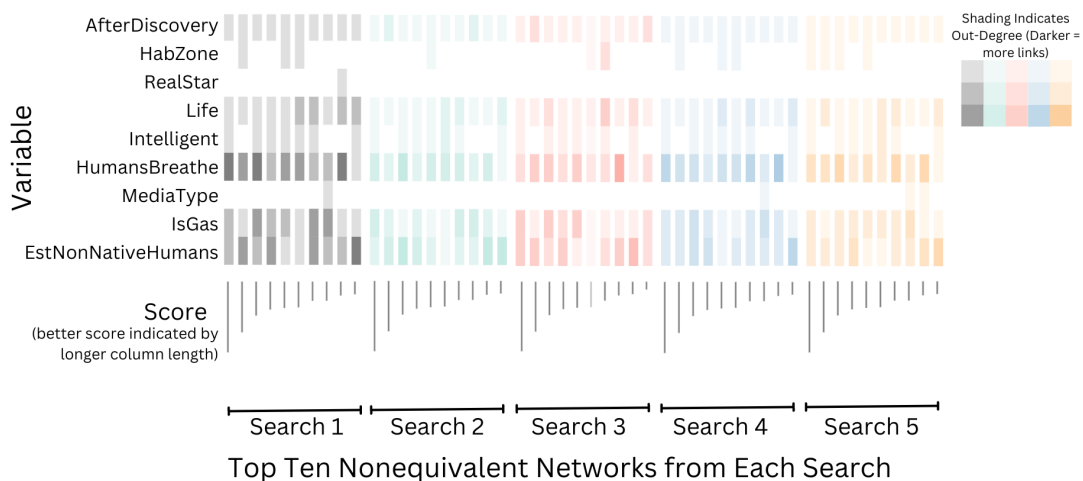
Figure 5.1: BayesPiles score-based visualisation of five greedy random-move 15-minute searches. Different searches are presented in sequential blocks. The top 10 nonequivalent networks from each search are shown, with scores increasing right-to-left indicated by line length. Network structure is shown in a collapsed form above each line by shading intensity indicating out-degree, or number of links, coming *from* each variable. Each search found top networks of the same equivalence class (mathematically equivalent, same BDe score) but the out-degree patterns are different.

The interface was also used to generate a matrix view of the network connections—the titular "pile" of BayesPiles. Figure 5.2 presents a "pile" of the five top-scoring networks—one from each search—all visualised in the same matrix. It can be seen that the darkest links are found in all five networks, but every link not found across all five networks in the same direction is still present in the opposite direction, i.e. EstNonNativeHumans–>HumansBreathe is found in one network and its reverse HumansBreathe–>EstNonNativeHumans is found in the other four networks. The consistency of the BayesPiles results clarifies how the Banjo algorithm built the networks and justifies the choice to present the fictional exoplanets Bayesian network as a single top network. In combination with the top network visualised in GraphViz (see next), the BayesPiles results illuminate the full context of the top network structure.
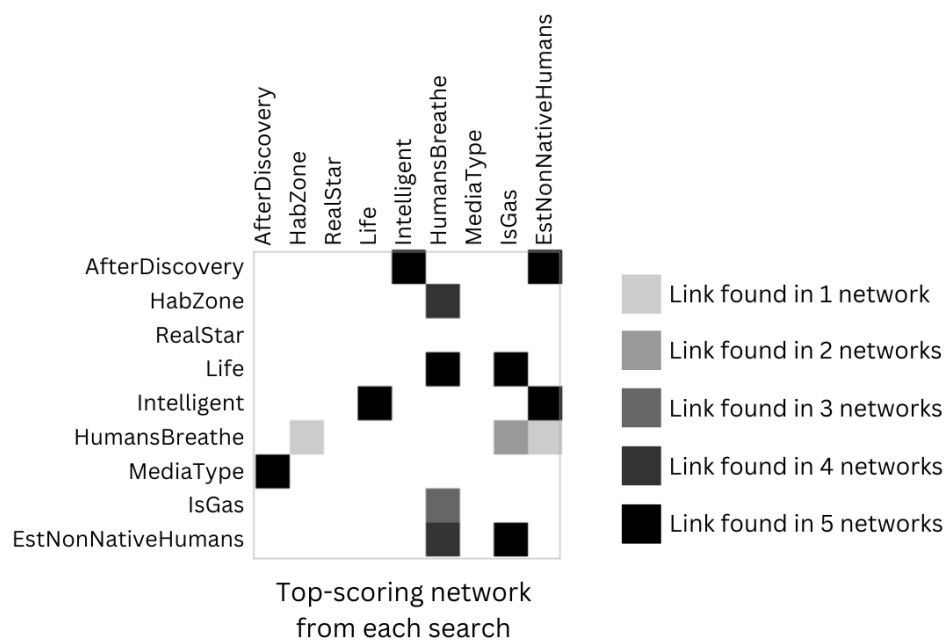
Matrix with Network Links from Column to Row

Figure 5.2: BayesPiles matrix-based visualisation of the top networks from each of the five greedy random-move 15-minute searches combined. Filled squares indicate links; the matrix can be read as showing links from the column variable to the row variable. The dark square in the middle of the top row for example shows a link from Intelligent to AfterDiscovery. Darker links are found in more networks; this representation allows the networks of the same equivalence class to become apparent. For each link found in 4 or fewer networks, there is another shaded square opposite it on a diagonal running from top left to bottom right; for example two searches found IsGas–>HumansBreathe and three found HumansBreathe–>IsGas; 2+3=5.

5.2   Top network

The highest-scoring network of the fictional exoplanet database is shown in Figure 5.3. Figure 5.3 shows the result of the first of the five greedy, random-local-moves, 15-minute searches performed. The labelled circles represent the variables, and the links between them show which variables were found to influence each other. A blue link signifies a positive influence; for example, if HabZone is true, it is more likely that HumansBreathe is also true. A red link signifies a negative influence; if HumansBreathe is true, it is less likely that IsGas is true, and vice versa. Full results including influence scores from all five searches can be found in Figure 5.4.

To summarise, the Bayesian network contains eleven links describing how fictional exoplanet characteristics influence each other. The five variables with strong positive links between them—Life, Intelligent, HabZone, EstNonNativeHumans, and HumansBreathe—form a cluster mediated by HumansBreathe. These features being true indicates a more Earth-like planet, showing the interconnectedness of these variables in science fiction worldbuilding considerations. There is a positive influence link between IsGas and EstNonNativeHumans, which will be further discussed below, as it is likely due to small sample size of gas giant planets and the effect of parent variables on influence scores (one run reported this link with a negative influence). There are negative influence links between IsGas and HumansBreathe and IsGas and Life, showing gaseous planets are negatively associated with these particular Earth-like features. There is also a negative influence link between EstNonNativeHumans and Intelligent, indicating that planets with intelligent native life are negatively associated with human colonisation efforts. There are also negative influence links between EstNonNativeHumans and AfterDiscovery, and between Intelligent and AfterDiscovery, showing that fictional planets with established non-native humans or with intelligent native life were less common after the real-life discovery of exoplanets. These negative links connect AfterDiscovery indirectly to two member variables of the positive cluster of Earth-like links, showing that of that cluster, EstNonNativeHumans and Intelligent are the two Earth-like traits likeliest to be absent in a post-detection fictional exoplanet. The network thus demonstrates that the changing likelihood of Earth-like fictional exoplanets before and after the discovery of real exoplanets is mediated by the presence or absence of intelligent life—be it native or in the form of non-native humans. There is a non-monotonic link between AfterDiscovery and our only nonbinary variable, Me-
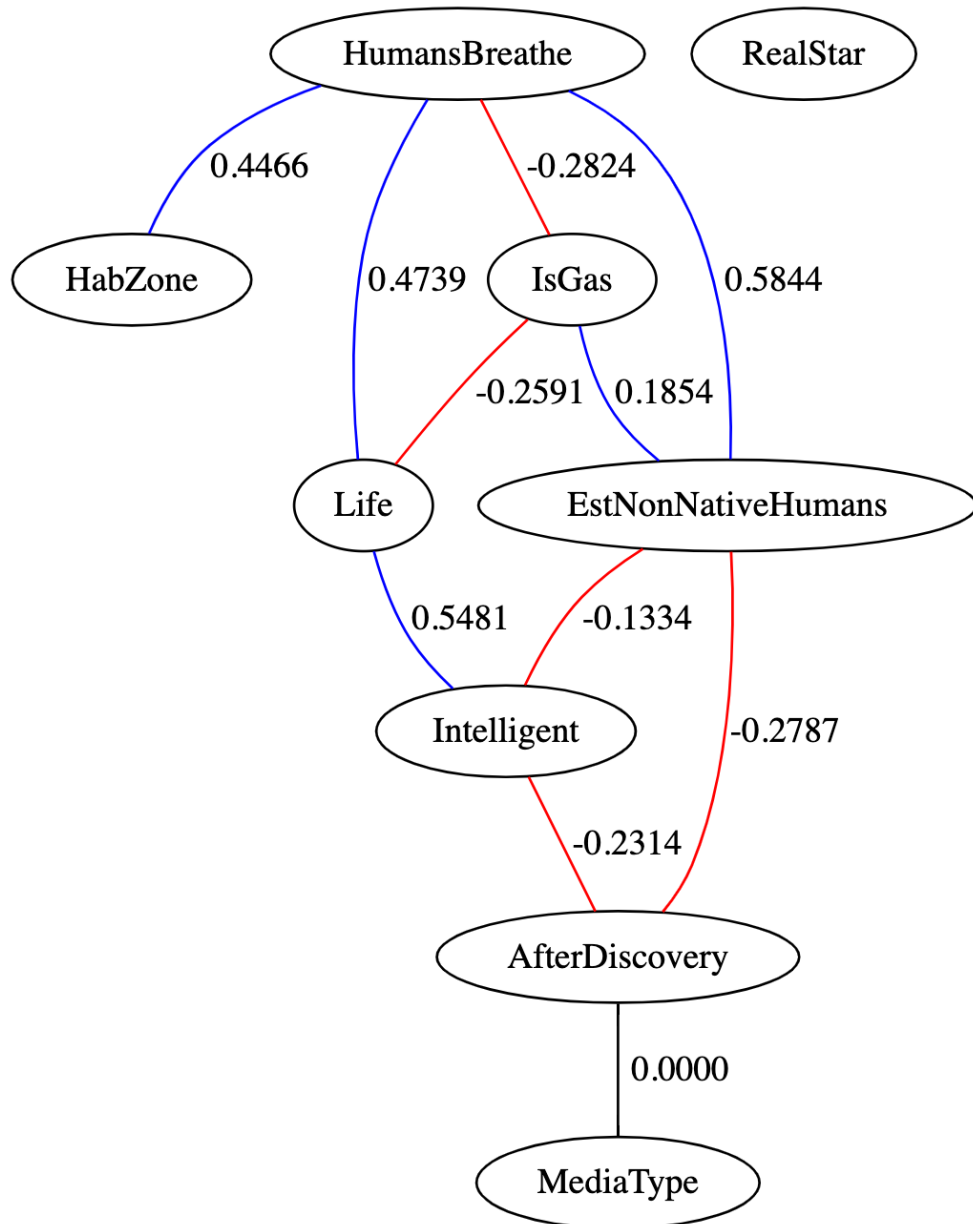
Figure 5.3: The highest-scoring network of the fictional exoplanet database. Positive links are blue, negative are red, and non-monotonic are black. Influence scores are next to links.

diaType, indicating there is an influence, but due to MediaType's having five states, it cannot be easily colour-coded as positive or negative in this visualisation. More recent SF stories are presented in a wider variety of media, and this link is caused by the increased prevalence of planets sourced from video games and podcasts in the post-1995 data. All variables except RealStar have at least one link to another variable, showing an absence of influence of whether or not the planet setting is in a real star system on other worldbuilding characteristics. The Bayesian network thus provides an overview of the landscape of science fiction exoplanets within my dataset.

	Search 1	Search 2	Search 3	Search 4	Search 5
EstNonNativeHumans→AfterDiscovery	-0.2787	-0.2787	-0.2787	-0.2787	-0.2787
Intelligent→AfterDiscovery	-0.2314	-0.2314	-0.2314	-0.2314	-0.2314
IsGas→Life	-0.2591	-0.2591	-0.2591	-0.2591	-0.2591
EstNonNativeHumans→Intelligent	-0.1334	-0.1334	-0.1334	-0.1334	-0.1334
Life→Intelligent	0.5481	0.5481	0.5481	0.5481	0.5481
HumansBreathe→Life	0.4739	0.4739	0.4739	0.4739	0.4739
IsGas→EstNonNativeHumans	0.1854	0.1854	-0.2490	0.1854	0.1854
AfterDiscovery→MediaType	0.0000	0.0000	0.0000	0.0000	0.0000
HumansBreathe→IsGas	-0.2824	X	X	-0.2824	-0.2824
IsGas→HumansBreathe	X	-0.6310	-0.2804	X	X
HumansBreathe→HabZone	0.4466	0.4466	0.4466	0.4466	X
HabZone→HumansBreathe	X	X	X	X	0.5862
HumansBreathe→EstNonNativeHumans	0.5844	0.5844	X	0.5844	0.5844
EstNonNativeHumans→HumansBreathe	X	X	0.5719	X	X

Figure 5.4: Influence scores of links found across five searches. Links in italics were found in both directions (ie from variable A to variable B in one search, from variable B to variable A in another) in different searches: if a link was found in one direction in one search, an "X" marks the opposite direction as it was not found in that particular search. Heavier line weights group together these cases.

5.3   Links

In this section, each feature of the final network will be discussed in detail.

Life—Intelligent, HumansBreathe—Life, HumansBreathe—EstNonNativeHumans, Humans-Breathe—HabZone

The network found positive links between many characteristics that, if true, indicate a planet may be more Earth-like. These are Life, Intelligent, HumansBreathe, EstNonNativeHumans, and HabZone. These links were found across all searches. Human intuition suggests that if a planet has life (with the caveat that this life bears resemblance to Earth life) then it is more likely humans can breathe there, and more trivially that if humans can breathe on a planet they are more likely to settle there. This group of positive links all have high influence scores, indicating strong relationships between these variables. The link between Life and Intelligent, in particular, is worth highlighting because Life must be true in order for Intelligent

to be true. The network reveals that Earth-like features positively inform other Earth-like features in fictional exoplanets, showing that authors often combine this set of characteristics in their science fiction worldbuilding.

IsGas—Life

IsGas and Life have a negative link. This finding is another that suggests that fictional exoplanets have very Earth-like features overall; Earth-like life is unlikely to be supported by a gas giant environment that has no solid surface. Despite a few exploratory papers on the possibility of floating life in a gas giant atmosphere (Sagan and Salpeter, 1976), and despite a few examples of gas giants with native life-forms in science fiction (Marijne VII in *Star Trek: The Next Generation,* Nasqueron in *The Algebraist)*, this result indicates SF exoplanets largely mirror reality in not presenting gas giant planets as homes for life.

HumansBreathe—IsGas

HumansBreathe and IsGas have a negative link. As shown in Figure 5.4, the top networks across the searches have different directionality but similar negative influence scores, except for search 2's result of a stronger negative influence score in the IsGas–>HumansBreathe direction. There is a small sample of fictional exoplanets in the database for which IsGas is true (see below on IsGas–EstNonNativeHumans), therefore the numerical results of the influence scores reported should be treated with less weight than the consistency of the negative influence result. Intuitively it makes sense that if human characters can breathe the atmosphere on a fictional exoplanet, it is less likely that that exoplanet is a gas giant.

EstNonNativeHumans—Intelligent

The network found a negative link between EstNonNativeHumans and Intelligent. The influence score is slight but consistent. This finding is interesting for suggesting that in science fiction, non-native humans are less likely to settle on planets that already have intelligent native life.

EstNonNativeHumans—AfterDiscovery

There is a negative link between EstNonNativeHumans and AfterDiscovery. In SF from after the discovery of real exoplanets, it is less likely for there to be established non-native human populations on exoplanets. This result is intriguing in light of real exoplanet discoveries being almost entirely of highly uninhabitable (by Earth life) planets.

Intelligent—AfterDiscovery

There is also a negative link between Intelligent and AfterDiscovery. Similarly to the above, this decrease in intelligent native life on fictional exoplanets worldbuilt after the discovery of real exoplanets can be read in the context of real exoplanet discoveries of planets very different from Earth.

IsGas—EstNonNativeHumans

The positive link between IsGas and EstNonNativeHumans is counterintuitive, compared to every other link the network found. Diving into the dataset, it is clear that IsGas being true is one of the least-represented types of fictional exoplanet, with only 24 out of the 212 planets in the fictional exoplanets database being gas giants. Of those 24, only two gas giants have established populations of non-native humans. These are Bespin, from the *Star Wars* franchise, and Goldblatt's World, from Larry Niven's *The Integral Trees*. For both of these planets, HumansBreathe is also true. Humans can breathe the atmospheres. As discussed in Chapter 3, influence score calculation is affected by parent variables; this is one of the weaknesses of the influence score (Yu et al., 2004). In this case, the link between IsGas and EstNonNative-Humans is found in the context of IsGas's parent variable HumansBreathe, making this a case where directionality *does* matter in order to understand what is happening. Because every time IsGas and HumansBreathe are both true for a fictional exoplanet (which only occurs twice), EstNonNativeHumans is also true, the Bayesian network has forged a link between IsGas and HumansBreathe with a positive influence score. This provides a useful example of how the networks are constructed, even though it is clear from considering the larger context that gas giant planets are not more likely to host established non-native human populations in science fiction. Notably, Figure 5.4 reveals that in the third of the five searches, a negative influence score was calculated between these same two variables. Search 3 is also the only search that found EstNonNativeHumans to be the parent in the EstNonNativeHumans-HumansBreathe link, thus changing the parents involved and resulting in the negative influence score. IsGas-EstNonNativeHumans is the only link for which this discrepancy exists; the rest are robust and found to either be always positive or always negative. In this case, the counterintuitive result is not due to a real trend but due to a small sample of gas giant planets.

AfterDiscovery—MediaType

AfterDiscovery—MediaType is the only non-monotonic link in the network, therefore its

influence score is reported as 0.00. This is due to MediaType having five categorical states (film, book, TV show, video game, and podcast), and the link found being therefore hump-or-U-shaped, between AfterDiscovery and various states of MediaType (Yu et al., 2004). In this case, while books, films, and television examples of fictional exoplanets were included from before and after the 1995 discovery of real exoplanets, the only video game and podcast fictional exoplanets in the database are from after 1995. (This is not to say neither existed before that date; as well, there are video games from the 1990s in the database, but they are all post-1995). The non-monotonic link between AfterDiscovery and MediaType is a result of changes in the media types in which SF stories are told over time.

Lack of links for RealStar

The variable RealStar has no links to any other variables in the network. This result has been consistent across every search (with the exception of the 9th-highest-scoring network in search 1, as shown in Figure 5.1), even in earlier research stages of this project in which the database contained fewer fictional exoplanets. RealStar is thus independent of the other variables in the network. Unlike IsGas, RealStar does have significant representation in the database, with 30% of fictional exoplanets orbiting real stars. It had been thought that RealStar might represent scientific knowledge of the author, and therefore perhaps negatively influence a variable containing a more scientifically improbable feature like HumansBreathe, but this is not the case. Whether or not the author has done the research to set their story in a real star system is independent of the other features represented by my variables.

5.4   Analysis

The research questions for the fictional exoplanets project were: *Exoplanet science has made massive discoveries in recent decades. Are these findings reflected in science fiction exoplanets, and if so, how?* and *Bayesian networks are graphical representations of interconnected variables. What interactions between portrayals of fictional exoplanets and scientific knowledge can a network reveal?*. The Bayesian network indicates that SF exoplanets are changing in light of new astronomical discoveries by becoming less Earth-like. The two most significant takeaways from the network are the strength of the Earth-like cluster, and the negative links between AfterDiscovery and Intelligent and AfterDiscovery and EstNonNativeHumans.

I am interested in what the network reveals about the portrayal of science in SF. Results

regarding AfterDiscovery represent what has changed in the process of imagining fictional exoplanets in the years since the real-life discovery of exoplanets. The network shows AfterDiscovery negatively influences Intelligent and EstNonNativeHumans, meaning that there have been fewer portrayals of intelligent native life and established non-native human populations on fictional exoplanets created since 1995. These negative links connect AfterDiscovery to the positive Earth-like cluster (HumansBreathe, HabZone, Life, Intelligent, and EstNonNativeHumans), showing a decrease in Earth-like characteristics post-1995.

Another research aim is to understand if exoplanet science findings are reflected in SF, and indeed, this decrease in Earth-like fictional exoplanets matches scientific findings. With the discovery of real exoplanets in 1995 and the findings of un-Earth-like classes of planets like hot Jupiters and tidally-locked Super-Earths, humanity has learned about a variety of exotic planets that are unlikely to be hospitable to humans. Popular news media reports on exoplanet discoveries, which would be accessible to SF authors, highlight hot Jupiters trailing comet tails of evaporating atmosphere (Choi, 2015), planets with temperatures high enough to rain diamonds (Strickland, 2022), and planets that orbit multiple stars (Carter, 2022). With increasing cultural awareness of extreme exoplanets and of the concept of a habitable zone from such reporting of exoplanet science findings, creators may have been inspired to explore what stranger new worlds outside this zone might be like. The Bayesian network reveals such changes in SF portrayal of exoplanet science with new discoveries.

Bayesian network analysis provides a statistical methodology with which to uncover trends in SF portrayal of exoplanets on a larger scale than traditional example-based close reading. The network can be used as an overview of the complete dataset, but it is also a starting place to look at detail in how science is portrayed in SF because it shows which characteristics inform each other. The network highlights that the AfterDiscovery links are mediated by the variables associated with intelligent life (Intelligent and EstNonNativeHumans). There are 67 planets in the database that have 1 for Life and 0 for Intelligent—that is, planets that host non-intelligent life. Of those planets that have non-intelligent native life but no intelligent native life, 69% of those from before the discovery of real exoplanets also have 1s for EstNonNativeHumans. Therefore, while intelligent life is depicted as living and thriving there (in the form of humans), it is intelligent life that is not native to the planet. Examples of unintelligent native life and established non-native humans include older works like Urras and Anarres from Ursula Le Guin's *The Dispossessed*, and Arrakis from Frank Herbert's *Dune*. This combination of char-

acteristics is not gone from more recent works—Nilt from Ann Leckie's *Ancillary Justice* is an example. However, planets with only non-intelligent life are more common in recent works: M6-117 from *The Chronicles of Riddick*, LV-1201 from *Aliens versus Predator 2*, and all four worlds from Becky Chambers' *To Be Taught, If Fortunate* serve as examples. The 69% figure can be compared to 34% among the planets with non-intelligent native life but no intelligent native life from *after* the discovery of real exoplanets. The Bayesian network thus illustrates, through its links between AfterDiscovery and Intelligent and EstNonNativeHumans, the story of how more recent science fiction is more likely to feature planets with un-intelligent native biospheres, and less likely to feature human colonisation efforts.

The strength of the Earth-like cluster in the network demonstrates the frequency of planets with Earth-like characteristics (life, rocky surface, human-breathable atmosphere, intelligent life) in SF. In investigating the portrayal of science in SF, this finding highlights that the population of fictional exoplanets looks very different from the population of currently-discovered real exoplanets. As discussed in Chapter 4, the dataset of known real exoplanets contains far fewer Earth-like worlds as a percentage than our fictional dataset, often in practice driven by narrative purposes in telling stories about humans. The fictional exoplanet population is much more Earth-like than the discovered exoplanet population, and there are still classes of real exoplanet I have found no representation of in science fiction. Hot Jupiters, for example, I can find none of in SF (aside from one example in the short story "The Stripped Core" in the *Around Distant Suns* anthology I edited, see Chapter 6). Goldblatt's World in *The Integral Trees*, a physically-impossible gas giant orbiting a neutron star and that has an atmosphere extended to the point of becoming a donut-shaped ring of sky around its star, is the closest other example I have found. The lack of Hot Jupiter representation in SF is another indicator that SF tends to centre human stories: Hot Jupiters are about the most inhospitable environments imaginable for human life. By contrast, real-life classes of exoplanet that writers can more plausibly worldbuild to be human-habitable, like ocean worlds and Super-Earths, *are* found in SF. As noted in the previous paragraph, however, more recent works are less likely to feature human colonisation while still featuring life-hosting planets, so perhaps a motivator for writers is not exploring the idea of humans in space, but exploring the idea of *life* in space. Applying Bayesian network analysis thus informs this addition to the discussion of the Earth-like natures of the fictional exoplanet population from Chapter 4, and demonstrates that habitable characteristics strongly influence each other in fictional exoplanets.

While SF exoplanets have an Earth-like bias, they still explore physical and astrobiological scientific concepts. Being set on a world where humans can breathe, or other scientifically-unlikely scenarios, does not preclude a work from sharing interesting and real scientific concepts, or from including un-Earth-like planets or un-human-like extra-terrestrials. SF includes the extrapolations authors make from scientific ideas. For example, the human characters can breathe the atmosphere on the planet January in Charlie Jane Anders' 2019 novel *The City in the Middle of the Night*. Yet, readers learn that January is a tidally locked planet—it rotates about its axis once per year, resulting in one side of the planet permanently facing its sun, and the other side existing in permanent darkness (Anders, 2019b). Tidal locking is a real physical mechanism that affects many planets, especially those orbiting close to their stars, as planets would have to around small and dim stars in order to be in their habitable zones. Astronomers think that about half the habitable zone planets discovered by the Kepler mission could be tidally locked (Barnes, 2017). In Anders' book, human settlers on January must contend with disturbed circadian rhythms due to the sky overhead never changing. Anders read science papers about tidally locked exoplanets and consulted with astronomers while writing her book, and reflects on the process: "Still, my tidally locked world didn't reflect these more recent computer models and ended up being a little more fanciful in some of the details. There's always a trade-off between scientific accuracy and storytelling, and in some ways, I may have ended up writing a bit of an exoplanet fable. But I wanted to help people imagine the strangeness, terror, and splendor of inhabiting a planet that orbits an alien star. I believe that novels about tidally locked worlds will become a fast-growing subgenre as we make more discoveries and gather more observational data" (Anders, 2019a). Here Anders explicitly references authors paying attention to new exoplanet science discoveries and featuring them in their works, in order to help the public imagine them. Even 'softer' SF franchises like *Star Wars*, which is considered emblematic of the space opera subgenre not renowned for its scientific accuracy, incorporate scientific concepts. Viewers see Tatooine and implicitly learn that there are binary star systems, and consider what it might be like to live on a planet orbiting one, as Luke Skywalker and his family do. Human characters can breathe on January and Tatooine, but these human-friendly worlds still introduce scientific concepts to the audience.

The Bayesian network allows us to see the big-picture, interconnected view of fictional exoplanet worldbuilding, and to make informed proposals for explanations for the links. For example, why are intelligent native life and established non-native humans negatively associ-

ated with fictional exoplanets from after 1995? My database indicates the majority of fictional exoplanets host native extra-terrestrial life. Alien environments and ecosystems might not interact well with human biological needs. Exoplanet discoveries, combined with a larger shift in understanding of environmental science and the rise of eco-anxiety (Coffey et al., 2021), have led to an increasing awareness of the interconnectedness of systems, and the ecological damage that could be wrought by separating biology from its own planet may be reflected in the changes shown by the network in recent SF. Human settlers on the titular planet from *Aurora* die of prion disease from their new exoplanet home (Robinson, 2015), and human settlers on Laconia in *The Expanse* cannot eat the native food on their new world because life there exhibits mirror chirality (Abraham and Woolnough, 2021). This new awareness may be behind the decrease in established non-native human populations on fictional exoplanets in recent works, just as increasing understanding of the variety of real exoplanet environments may be behind the decrease in exoplanets with intelligent native life.

Taken in combination, my findings of the Earth-like cluster and its negative links with AfterDiscovery, as well as the lack of links involving RealStar, suggest that an author/creator's level of scientific knowledge is not reflected directly in the features of exoplanets imagined. Instead, these features are more connected to the gestalt presented by the overall community of knowledge, a big-picture finding made possible by the Bayesian network methodology. That fictional exoplanets are becoming less human-habitable is a demonstration of exoplanet science findings within science fiction, with the Bayesian network revealing the interconnected variables at play.

Conclusions should be seen in the light of the limitations of the study, and there is a strong potential for future work to be done. Chapter 4 discusses the need to only include fictional planets for which all variables could be answered, due to challenges with missing data in Bayesian networks. Because of this, variables were selected to be as widely-applicable as possible. This investigation into exoplanets in science fiction represents a preliminary evidencing of the role of science in SF, and future work could explore delineations between hard and soft SF, or SF that features terraforming (the process of altering a planetary environment to be more Earth-like, for example).

Bayesian network analysis has the power to uncover larger trends within the genre of science fiction, and from other digital humanities and media studies datasets, that are impossible

to learn from close readings alone. This methodology is highly applicable to different problem sets. Trends learned from Bayesian network analysis could be used to guide investigations into the corpus of science fiction media for purposes of teaching science, by helping select fictional exoplanets with characteristics of interest for teaching. While my findings point towards a strong potential for a role for SF in science education and provide an adaptable methodology for monitoring science content in SF, further research could develop methodologies for studying hard versus soft SF, or developing curricula to teach science using SF. Through the application of Bayesian network analysis, I reveal changes in SF media towards less Earth-like and human-habitable portrayals of fictional exoplanets. Results show that science fiction as a genre does evolve alongside scientific discovery, reflecting the findings of the growing field of exoplanet science.

6

Centre for Exoplanet Science anthology project: an inspirational, exploratory role for science in SF

"Fantasy is totally wide open; all you really have to do is follow the rules you've set. But if you're writing about science, you have to first learn what you're writing about."

—Octavia Butler

[Portions of this chapter are taken from my paper "Dialogues Between Science and Fiction in the Creation of an Anthology", published in BSFA Vector 297, special issue on applied speculation]

Science fiction reflects what its writers see in the world around them—often from current scientific discoveries—and it sparks ideas for scientists. As discussed in Chapter 2 with the "revolving door" analogy, scientists and SF writers endlessly inspire each other, with scientists making discoveries, writers extrapolating on them, and scientists growing curious about how

such extrapolations might work in reality. However, little research has been done on how exactly this inspiration happens—on the dialogues and interactions between the two often-overlapping groups of SF writers and scientists. Given SF's reputation for applied speculation and future thinking discussed in Chapter 2, these dialogues are key to any studies of the same, and to those interested in determining the uses of SF for science education. I address this gap through analysing qualitative data on the experiences of scientist and writer participants in an SF anthology project which included significant interdisciplinary encounters.

Around Distant Suns: Nine Stories Inspired by Research from the St Andrews Centre for Exoplanet Science (2021) is an SF anthology that I conceived and edited, containing five short stories, two radio play scripts, and two poems. Each contribution was created by a pair of one scientist and one writer, and has a basis in the scientist's research. The St Andrews Centre for Exoplanet Science (https://exoplanets.wp.st-andrews.ac.uk/), founded in 2017, produces research addressing questions that form some of the core themes of SF, therefore it provided a source for scientists for the research project. My goal was to investigate how scientists and SF writers work together in creating science fiction stories, with a particular focus on the processes of deciding when to stay realistic, when to be plausible, and when to make things up—in short, to determine how science fiction worlds are built. This research was undertaken independently of the fictional exoplanets Bayesian network.

I present results from qualitative analysis of the questionnaires, which asked about communication successes and failures, challenges encountered and solved, and when and how story decisions were made and inspired. Results point to a role for SF in science education of introducing concepts and piquing curiosity, but, in keeping with Suvin's idea of estranging the worldviews of the readers, also leaves room for the fantastic and the unknown.

6.1   Background

As discussed in Chapter 2, the genre of science fiction has a unique relationship with empiricism in its worldbuilding, which I argue results in the mutual inspiration between science and SF, and the genre's reputation for being at the forefront of scientific discovery. Sources of scientific inspiration and the degree of superficiality or robustness of the fictional science are as varied as the genre itself. In addition to scientists writing SF, and workshops like NASA's Launchpad that offer writers the chance to liaise with scientists in the creation of their works, there are
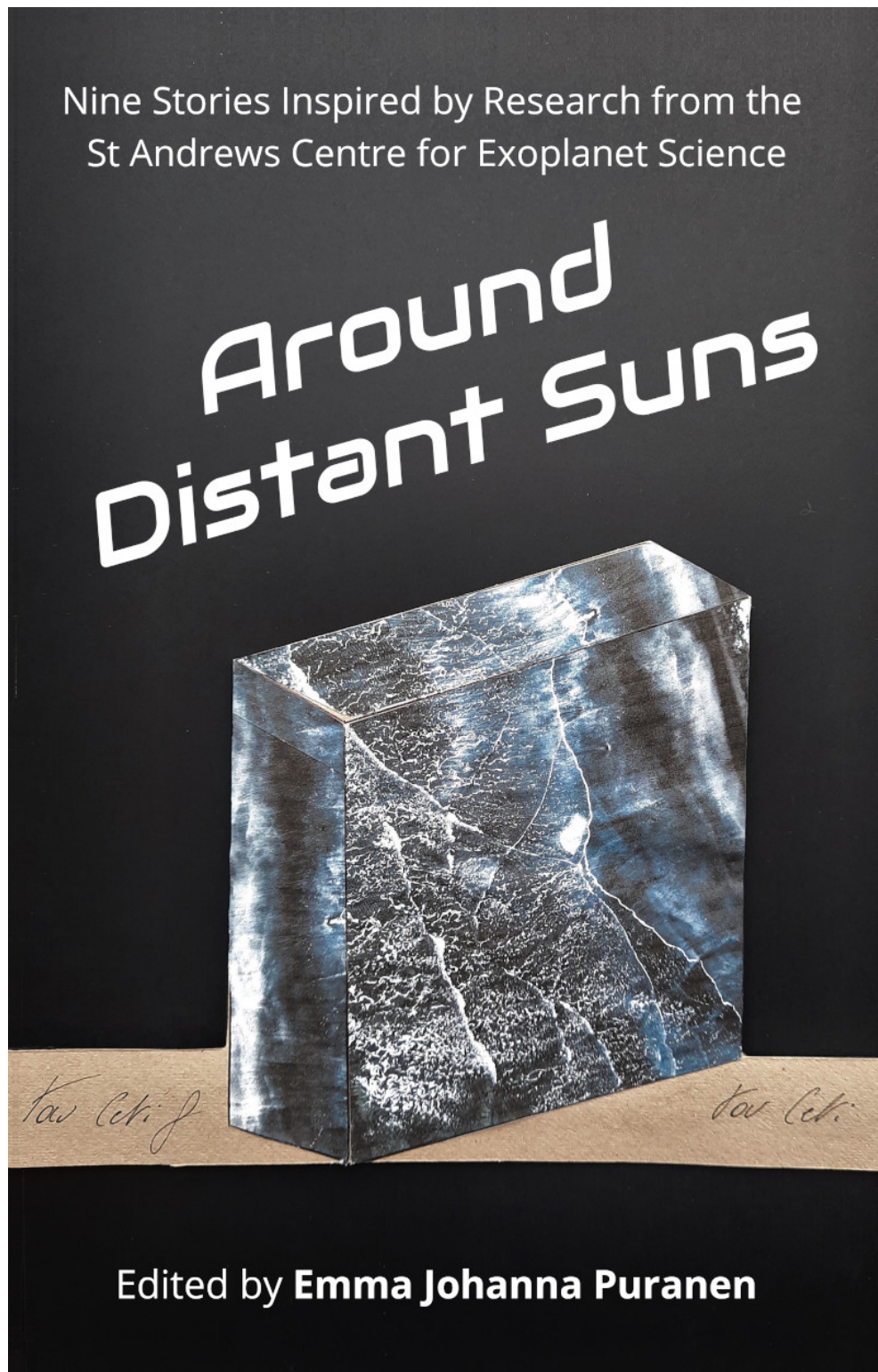
Figure 6.1: Cover for *Around Distant Suns*. Art by Tinne de Vis.

many other ways in which *science consultants* can be involved in the worldbuilding process for SF stories. Acknowledgements sections of SF novels are often filled with references to e-mail exchanges with science consultants. The science consultant is key for understanding decision-making processes regarding portrayal of science in SF.

Science consultants can have many different levels of involvement in the story process. Sometimes, it is a single, quick discussion. Physicist Sean Carroll, science consultant on several Marvel movies including *Thor*, describes "You talk to the screenwriter or director or producer—whoever asked for your help—and you chat for a couple hours, and you do your best to give them advice, and then you never hear from them again" until the film has its theatrical release ("Science Consultant", 2019). Occasionally, a scientist has a heavier involvement, and sometimes scientific publications can result. Physicist Kip Thorne, an expert on black holes, is one such example. Carl Sagan, in writing the SF novel *Contact* in the 1980s, asked Thorne for help in how his lead character might be transported a great distance in a short time; Thorne suggested using the concept of wormholes, which he had written papers on (Sagan was a planetary scientist and therefore asking for assistance outside his field of study) (Carroll, 2013). Thorne later went on to famously make real scientific advances in determining the optical-wavelength appearance of a black hole for the 2014 film *Interstellar* (James et al., 2015). When the first photograph of a real black hole was taken in 2019, Thorne's simulated one from *Interstellar* was proved quite accurate—see Figure 6.2 and Figure 6.3 (Akiyama et al., 2019). Thorne provides an example of how being a science consultant can inspire scientists in their careers through the questions prompted, rather than the consultant simply being a one-way flow of information. However, unless the writer themself is also the science consultant, science consultants rarely play an equal role in story creation.

Curious about the role science consultants play in the creation of SF and the presentation of science therein, I set out to produce an anthology creating science-based SF in which science consultants remained active participants throughout the entire project, in order to understand the role of science in the decision-making process in SF worldbuilding. My research questions were as follows:

Research Questions: What is the role of a science consultant for a SF story? How are decisions made regarding the portrayal of science in SF? What is the role of SF in science education?
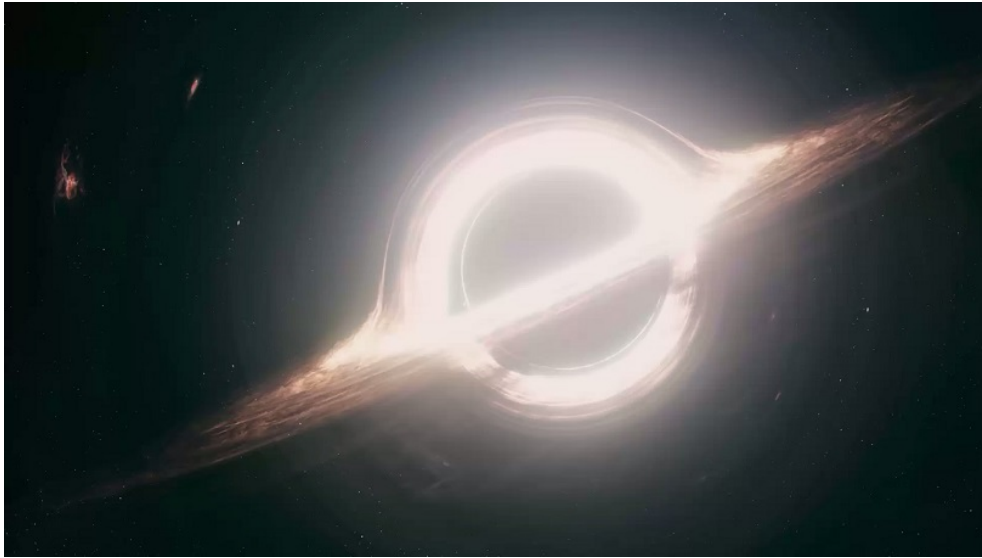
Figure 6.2: The black hole "Gargantua" in the film *Interstellar*, appearance based on research by Kip Thorne. Credit: Paramount.

Figure 6.3: The first image of a black hole, providing observational evidence to support Thorne's theoretical work. Credit: Event Horizon Telescope Collaboration.

6.2   Results

The full methodology is described in Chapter 3. Five teams participated in the research portion of the anthology. They are pseudonymised in the following manner: 1S and 1W refer to the scientist and writer from the same team, 2S and 2W are from the same team, etc. Every participant submitted at least one questionnaire, with the total number of completed questionnaires from each Team varying from three for Team 2 to 17 for Team 5.

The Questionnaire asked:

What was discussed at this meeting?

What story ideas were generated?

What story decisions were made?

Describe any communication difficulties.

Describe any communication successes.

Describe the current status/progress of the story.

Results are qualitative, and were collated for common occurrence of themes. Participants wrote as much or as little as they wanted in answer to each question, and did not have to answer every question. All teams succeeded in creating a story—two teams wrote scripts for radio plays/audio dramas, and three opted for short stories. Below, I list elements that recurred in at least two of the five teams, with the numbers of the teams each element applied to following in parentheses. Notably, a team not being listed does not mean this element was not present in their meetings and creative process, but only that it was not recorded on the questionnaires:

Hesitance of the Scientist: Despite the instruction that the scientist remain involved throughout the process, scientists often expressed hesitation to contribute plot ideas authoritatively due to lack of experience, preferring to leave those to the writer. (1, 2, 4)

Scientist provides justifications: The writer wanted a certain setting for the story or event to take place, and the scientist provided a scientific justification for that setting or event. (1, 3, 5)

Scientist provides technical terminology: Scientists from every team provided accurate tech-

nical terminology from their subfield for the writer to use. (1, 2, 3, 4, 5)

Research work undertaken during writing process: Either the writer or the scientist, or both, consulted scientific sources for the story creation process, sometimes as 'homework' between meetings. (1, 3, 5)

Focus on fieldwork and roles of scientists: Writer expressed strong interest in hearing about and incorporating scientist's fieldwork experiences. (1, 3, 4)

Discussion of shared interest in SF: Often at the first meeting and as a way to form an initial connection, the team discussed SF they are fans of, and what they like in an SF story. (1, 4, 5)

Scientist-focused Question-and-Answer at first meeting: The first team meeting was largely a question-and-answer session with the writer asking and the scientist answering. (1, 2, 3)

Balanced Question-and-Answer at first meeting: The first team meeting was a question-and-answer session where writer and scientist each spent time questioning and answering. (4, 5)

Focus on sense of ambiguity/mystery: The team discussed the big mysteries of science and chose to incorporate a sense of mystery or the unknown into the story. (3, 4, 5)

6.3 Discussion

Teams had diverse experiences of the collaborative creative process. At the beginning of the process, two different models were enacted for the running of the first meeting: in the *Scientist-focused Question-and-Answer* groups (as designated by my research finding), information largely flowed from the scientist to the writer. In the *Balanced Question-and-Answer* groups, both participants described their backgrounds and asked questions about the other person's background, with Team 5 in particular describing 5W sharing their story-crafting techniques to 5S. At the first meetings, contributors often expressed a lack of surety. 1W said they "hadn't done much planning for the story beforehand because I was unfamiliar with [1S's] research and knowledge area, so it felt a little like we were both stumbling around in the dark not sure where to begin." 3W described "I think I am still processing a lot of what [3S] said. It felt a bit like being swept up by a wave—exciting, but my feet were definitely not on the ground. There are gaps in my understanding and I'm waiting to see if these need to be addressed." These statements came from writers of the Scientist-focused Question-and-Answer

groups, with the Balanced Question-and-Answer groups writing about how, as 4S described "We had to 'find each other' at first, to figure out how much each of us knows about the other person's field. There were a few terms one or the other didn't know at first, but these were simply explained, and personally I enjoy learning something new!". Discussion of shared experiences with SF often provided a natural point of commonality and jumping-off point for these initial conversations, for example, 1W referenced an *X-Files* episode as an influence for the plot of their Team's story. Team 5 also bonded over shared interests; 5S wrote "Both showed great interest in the others field of study. We strongly connected over personal interests in the others field of expertise ([5W] likes to read about space/astrophysics; I like to write tabletop screenplays and poems)" (*Discussion of shared interest in SF*). Many contributors wrote that they were concerned they were not communicating their ideas well, yet every contributor stated they felt their teammate had communicated well. Therefore, despite much concern about communication troubles, no major miscommunications were actually recorded. 1S stated "We quickly found some common ground, e.g., we are both very interested in exploration as a general theme, and in scientists as interesting characters."

Several Teams left their first meeting planning to undertake research in the interim before their next meeting, with Team 1 for example agreeing to each look at the NASA exoplanet database before the next session (*Research work undertaken during writing process*). Team 5 was the only Team to assign writing "homework" in addition to science research, with both members working on an exercise about finding character set out by 5W.

All Teams had to determine the role of the scientist in the writing process. There were a variety of solutions to this, with some scientists largely providing information and reading drafts, while others commented on plot and character decisions. 5S was the most involved scientist, completing writing exercises and helping build a story arc (Team 5 met nine times). Several scientists displayed *Hesitance of the Scientist* to contribute to plot. 1S wrote "I'm definitely not a creative writer so very happy with my role being largely inspirational, rather than contributing to specific plot development...I am trusting [1W] to make specific story decisions, but am enjoying contributing broad ideas." 2S wrote "I left most of the specifics of story decisions up to [2W]" aside from stating what they did not want to include in the plot, such as extra-terrestrials. While 4W wrote "I was responsible for story boarding something out of this and [4S] for chipping in with ideas if [they] wanted and as requested so we can tie the story as close as possible to [their] research", 4S confirmed "All large decisions are ultimately taken

by [4W]". No scientists actively wrote parts of the final stories, with Team 5 deciding "We struggled to find a way to work together on the [story] skeleton but decided that [5W] will first write an outline and I add my comments" (5S).

All scientists provided accurate technical terminology and scientific facts that were incorporated into the stories (*Scientist provides justifications, Scientist provides technical terminology*). Occasionally some science points were insisted upon, such as locational accuracy when mentioning a real star system, or locations on Earth most suitable for astronomical observation. 1S advised on the differences between an atmosphere in which human characters could breathe versus what it would mean if the main character were wearing a spacesuit, and 2S provided real exoplanet names for the story. 5S undertook research between sessions to inform plot points. However, many scientists wrote of feeling the pressure to communicate their science accurately, and as a result were careful to differentiate their own views from the prevailing views in their field, and to emphasise the unknowns of science. For instance, 2S did not want to include extra-terrestrials in their story, as they felt it would constitute heavy speculation for their subfield; 3S "discussed the insufficiency of claiming an authoritative interpretation on the basis of the current state of knowledge", and 5S emphasised the need to admire the great mysteries of science, not just answer them.

In addition to providing scientific information, scientists also advised on everyday life as a scientist and how scientists are seen by society, and the purpose of science (*Focus on fieldwork and the roles of scientists*). 4S wrote "We also discussed the roles of scientists in the society, and how difficult it can sometimes be for non-scientists to understand the scientific results." It was in these discussions that writers sought out the emotional weight in the scientist's work. 1W "enjoyed getting to hear about [1S]'s experiences in the field...it helped build the emotional truth of field research for me, which is really what stories are trying to capture and convey while technical details are just the icing on the cake" and consequently decided to make the environment key to the plot of their story. There was a lot of interest from writers in hearing about field work experiences, possibly because it is easier to tap into story themes and emotional weight when writing about characters in the field rather than sitting behind computers. 5W added a "seed of loneliness in the story, because one of the reasons we look at exoplanets is to see if we are alone", explicitly addressing the big questions of exoplanet science. Such themes often prompted scientists to bring up new scientific concepts they felt were relevant—5S brought up the death of stars and the cycle of matter in response to the

loneliness idea, which was then incorporated in the story in a collaborative process. This exchange provided an example of the back-and-forths that happen in relationships between SF authors and scientists—rather than a single instance of inspiration from science to the author, it is instead a discussion that each contributor adds to multiple times.

Teams reflected throughout the process on the role of science in their stories, and the potential for educative uses. There was a common desire to use the stories to inspire readers into beginning their own investigations into science, rather than simply laying out facts. Team 2 reflected "in general the scope of the story was good, and although relatively light on the scientific detail, serves the purpose set out in the first meeting to rather try and inspire an interest in non-science based readers and rather not get bogged in the details of the science" (2S), and stated they hoped including scientific names and jargon would function as a "'jumping off point' for readers to look up some real science". S3 similarly had a goal of "education not being about telling things, but enabling to explore and arrive at conclusions." The stories were seen as a chance to inspire, rather than to teach. This conclusion highlights a key difference between science and fiction: science seeks answers, whereas fiction sometimes functions best leaving some questions unanswered—and science fiction faces a particular challenge of using narrative devices to communicate unknown speculative science to its readers (Slusser and Chatelain, 2002). 4W highlighted this when discussing a decision to not provide a reason for an action, writing "We then decided it would be more interesting to leave this ambiguous, so an audience could fill this in for themselves!" (*Focus on sense of ambiguity/mystery*). As 5S expressed a belief that science is sometimes about "admiring the great mysteries", 5W wrote "I think a great story is all about a question, but it's not up to the writer to answer it, it's for them to explore." Thus the science in SF allows for not only inspiration, but exploration of scientific possibilities.

All contributors expressed that they enjoyed the collaboration — scientists enjoyed sharing their work, and writers enjoyed sharing what is typically a lonely process. As 5W put it "I really enjoy collaborating in writing; it's a medium that's often unnecessarily gatekept". 1S, in addition to teaching 1W about astrobiology, was excited to have learned more about writing and poetry. 1W found having someone there for "helping me brainstorm ideas to help me get through my writer's block and figure out what should happen in the middle of the story" helpful. 3S found the experience helpful as a scientist, writing "I felt inspired through the interaction, and in fact new thoughts were triggered from it". 5S found the collaboration to

flow both ways, writing "I can give [5W] foundations to [their] story ideas which makes them more physical and improve the story. [5W] can adjust my inputs to better fit a good narrative" and that their interdisciplinary collaboration was successful because "We both are comfortable to reach into the others field of expertise with the expectation to be corrected and learn from it."

6.4 Conclusions

This sort of interdisciplinary collaboration was new to many of the participants. A number of them were familiar with interdisciplinary conversations, but less so with a collaboration that produced a piece of writing. Challenges included scientist hesitance to contribute to plot, and occasional one-way flows of information from scientist to writer. Every Team found a different role for the scientist, varying in involvement in plot decisions, showing there are many active ways to be a science consultant for an SF story. Teams largely moved from feeling somewhat overwhelmed in communicating with each other, especially regarding the unknowns, to growing more comfortable with these unknowns, to very often incorporating them into the final stories. Decisions were made collaboratively, with a desire for scientific accuracy often informing plot points. It was commonly concluded that the role of the science in the stories was to provide jumping-off points, or inspiration, to readers, as well as a chance to explore scientific questions. Writers were eager to include some amount of real scientific jargon, and there were certain science elements scientists insisted upon, but ultimately some room was left for the fantastic and the unknown. This is in keeping with Suvin's idea of estranging the worldviews of the readers. Thus, participants carved out this role—inspirational, exploratory, not very strictly constrained by accuracy, introducing concepts—for SF in teaching science.

Conclusions should be seen in the context of the uniqueness of this collaboration, as well as the small sample size of Teams, and not generalised beyond these circumstances. The *Around Distant Suns* anthology project was a planned and structured collaboration taking place in an interdisciplinary university environment with an explicit research component, a rare set of conditions for the writing of science fiction; as well, the scientists and writers were on more equal footing and closer to being peers than in many SF creations involving a science collaborator, in which, for example, the writer may be a famous figure, thus this experimental set-up did not establish hierarchies. However, common takeaways regarding the role of the unknown in science and in SF, and the inspirational value of SF in introducing science concepts

to larger audiences, are valuable to those interested in the intertwining of the two fields.

7

The N=1 Problem in exoplanet science

"It was strange to have sailed the sea for many years in entire ignorance that such things were."

—Robert Falcon Scott, *Voyage of the Discovery*

"The only thing that makes life possible is permanent, intolerable uncertainty: not knowing what comes next."

—Ursula Le Guin, *The Left Hand of Darkness*

[Portions of this chapter are taken from my chapter 'Through a Mirror, Earthly: Solaris, Gaia, and the Search for Other Worlds' in the collection Stanisław Lem and His Aliens: A Tribute and a Challenge (2022)]

7.1 Motivation

Central to all scientific inquiry regarding life in the universe is the *N=1 Problem*, the fact that humanity only knows of one planet with life on it: the Earth. How do we search for life in the universe from a sample size of one? Can we determine which qualities of Earth life and Earth's environment are universal to life, and which are unique? When exoplanet science emerged as a significant subfield of astronomy after the discovery of the first exoplanets in the 1990s, astronomers found themselves involved with the question of life beyond Earth moreso than ever before. While the scientific search for extra-terrestrial intelligence, or SETI, had been ongoing for decades already using radio telescopes to "listen" for signals from space, scientific inquiry about exoplanets, as discussed in Chapter 2, had always been a more theoretical effort that lacked data. At the time of writing, 5599 exoplanets have been confirmed on the NASA Exoplanet Archive (NASA Exoplanet Archive, 2024) and 5652 on exoplanet.eu (exoplanet.eu, 2024), resulting in large quantities of observational data about these planets now being available. Exoplanets are one of the top three research priorities for upcoming large missions in the European Space Agency's Voyage 2050 paper (another is icy moons in the solar system, which also represent a research area of strong astrobiological interest (Calapez et al., 2023)) (ESA Voyage 2050, 2021). Exoplanets are also one of the top three science focuses in the National Aeronautics and Space Administration's 2020 Decadal Survey, with "Pathways to Habitable Worlds" listed as the priority area of the exoplanet science focus (National Academies Decadal Survey, 2021). A strong influence on why exoplanets are of interest to these international space agencies is addressing the question of whether there is or could be life elsewhere in the universe—this is implicit in the use of the word "habitable" in both survey papers, and NASA's decadal survey explicitly references the "Are we alone?" question (National Academies Decadal Survey, 2021). Astronomers are not trained biologists and astrophysics cannot answer questions about the nature of life; but due to the relevance of exoplanet science to astrobiology—the study of life in the universe—astronomers are collecting data that is *of interest* in answering questions about biology. This motivates my investigation into the effect of the N=1 Problem on the field of exoplanet science.

This chapter presents results from a study of 18 researchers in the field of exoplanet science asked about the significance of the N=1 Problem on their research. This study is integral to my wider research questions about the relationship between science and SF, and about SF

use in science education, because the N=1 Problem affects SF portrayals of extra-terrestrial life and exoplanets, as well as being a reality of exoplanet science work. Participants were interviewed about their familiarity with the N=1 Problem, its effects (or lack thereof) on their work, and its effects on the field as a whole. The full methodology is given in Chapter 3. Interview results included a wide variety of responses to the problem, with the two main research responses being: attempts to be less Earth-centric in research methodology to avoid missing or failing to recognise life, or using Earth's example to guide research methodology as it is the only empirical evidence present. Results indicate that while all participants were familiar with the N=1 Problem, none were formally taught it, and it is often discussed outside the strict context of their academic publications. Techniques for working within the N=1 Problem suggested by participants included increased interdisciplinarity with astrobiologists and improved interdisciplinary funding and publication infrastructure to support this, as well as better communication of the uncertainties in science to the public. This study is preliminary, and its goal is to begin investigating the range of thoughts, opinions, and areas of opportunity, concern, or challenge as seen by current practitioners in the field. Importantly, results should be considered in the context of the small sample size of researchers interviewed and a possible selection bias of individuals already aware of the challenges of the N=1 Problem.

7.2 Background on the N=1 Problem

It is necessary to define and differentiate exoplanet science and astrobiology for the purposes of this chapter. Both are relatively new, highly interdisciplinary fields, and though they often collaborate together they have different objects of study, exoplanet science originated in astronomy (see Chapter 1) while astrobiology originated in biology (McMahon, 2021). *Exoplanet science* is the scientific study of exoplanets, or planets outside our solar system, and has evolved from the original detection of such planets in the 1990s to include theoretical modelling of their atmospheres and compositions and large survey and characterisation space telescope missions. It is often considered a sub-field of astronomy, the field of science that made the initial discovery of exoplanets, although it has evolved to become a "poster child for multi-disciplinary collaborative science" incorporating aspects of geology, biology, and planetary science, for example (Howell, 2020). Exoplanet science covers all exoplanets, regardless of how similar to Earth they may be, and whether or not they are considered plausible candidates to host life. *Astrobiology* is the study of life in the universe. It aims to study life as

a universal phenomenon, encompassing research on the origin of life on Earth, on the conditions for life, the potential and search for life in the universe, and the future of life (Dunér et al., 2018). According to NASA historian Steven J. Dick, astrobiology emerged in part from the narrower field of exobiology, which was focused solely on life detection (Dick, 2020b). Sometimes the field is, often jokingly or by detractors, referred to as a field without a subject or seen as less legitimate science due to the fact that as yet we know of no confirmed life on planets other than Earth (the N=1 Problem) (Billings, 2012). This, however, is a misrepresentation of astrobiology's substantial contributions to science in the form of extremophile research, investigations into Earth's history, and the characterisation of environments such as Mars and sub-surface oceans of icy moons—astrobiology is not just a search for life, but an extrapolation of the field of biology to a universal scale (McMahon, 2021). Astrobiologists tend to come from a more diverse range of educational training than exoplanet scientists, including scientists with degrees in biology, earth sciences like geology, and chemistry, although a few specific degree programmes exist (UW Astrobiology, 2024). Both exoplanet science and astrobiology have emerged in recent decades to see increases in research output, influence, recognition, and funding.

The N=1 Problem is a current consideration at the intersection of exoplanet science and astrobiology that this study seeks to investigate. To clarify, I am using the term "problem" in the scientific sense, in that this is a question that can be investigated through evidence-gathering and experimentation. "Problem" does not connote a value judgement or imply that the N=1 Problem is negative or an issue, in the more colloquial sense of the word. Rather, it is the current state of the study of life in the universe, and it affects to varying degrees efforts in the field of exoplanet science, as well as the ways humans envision exoplanets in general, including in SF. It is for this effect on exoplanet science that work on the N=1 Problem is important for my overall research. The motivation for this project is to investigate how current exoplanet scientists work within, around, and alongside the N=1 Problem.

The N=1 Problem is defined as the fact that we only have one example of a planet with life—Earth—and consequently there may be doubt that principles of Earth biology would translate to life elsewhere (Mariscal, 2015). According to Mariscal, the N=1 Problem is, within the scope of human history, a relatively recent phenomenon. Before Darwin set off the series of biological understandings that let to the knowledge that all Earth life is closely related and probably has the same last universal common ancestor (LUCA), it could plausibly have been

thought that Earth life has multiple origins. This would solve the N=1 Problem in the sense that we would have multiple examples of life of different origins on Earth alone. Since all Earth life has the same origin, we only know of one example of life, and we have the N=1 Problem: how can we learn about universal biology from a single example of life? Carol Cleland writes "the most significant challenge facing the pursuit of a universal theory of life is in the infamous 'N=1 problem' (Cleland, 2019). The N=1 Problem is a central consideration in astrobiology as it effects the study of the universality of life.

The definition of life, in fact, remains a question today. Scientists do not have one clear definition for life; different fields have different definitions, all of which capture some characteristics but none of which are fully satisfying. The autopoietic (referring to a system that can reproduce itself) definition of life holds

- "an autopoietic system—the minimal living organization . . . [is] a network of processes of production (synthesis and destruction) of components such that these components: (i) continuously regenerate and realize the network that produces them, and (ii) constitute the system as a distinguishable unit in the domain in which they exist" (Varela, 1997).

The Darwinian definition, used by NASA as a working definition, says that life is

- "a self-sustained chemical system capable of undergoing Darwinian evolution" (Brennan, 2023).

The autopoietic definition is non-concrete, and the Darwinian one suffers from its inability to account for animals that cannot reproduce, such as mules, and our lack of certainty about the universality of Darwinian evolution among life (Cortesão, 2015).

Carl Sagan defined separate physiological, metabolic, biochemical, genetic, and thermodynamic definitions for life, noting "Man tends to define in terms of the familiar" (Sagan, 2010). Writing in 1970, Sagan addressed the N=1 Problem in his summary: "It is difficult to generalize from a single example, and in this respect the biologist is fundamentally handicapped as compared, say, to the chemist, or physicist, or geologist, or meteorologist, who now can study aspects of his discipline beyond the Earth" (Sagan, 2010). While such varied working definitions of life are useful, each definition again is lacking in some way, with biases towards certain characteristics and the priorities of certain scientific subjects (Cortesão, 2015). The

field of astrobiology faces a challenge in that it may well be impossible to define life until we find non-Earth-based life, yet without a definition for life, we may not be certain we can recognise such non-Earth-based life. The challenge comes from the simple fact that the only life we know of is from Earth, and drawing conclusions from a sample size of one tends to lead to biased answers.

Given the N=1 Problem, life detection research is a challenge: one recent example of this is the case of phosphine on Venus. This research was in the field of biosignature detection. Biosignatures are signs that might indicate the presence of life—for example, certain gases present in a planet's atmosphere—but could also have abiotic sources (Schwieterman et al., 2018). Although evidence gathered from the first Venus missions in the 1960s and 1970s pointed towards an inhospitable pressure cooker of a planet inhospitable to life as we know it, the search for microbial life on Venus is still an active area of research. Although the surface experiences extremely high temperatures and pressures, high up in the cloud decks, there is thought to be an environment more amenable to life. Communities of microbes in the clouds of Venus have been theorised for decades (Morowitz and Sagan, 1967).

In late 2020, Greaves et al. published a paper in *Nature* announcing the discovery of phosphine in the atmosphere of Venus. Phosphine is considered a strong biosignature. On Earth, it is uniquely associated with biological origins, and phosphine should quickly be destroyed in the atmosphere if it is not continuously replenished (Greaves et al., 2021). Greaves et al. reported a detection of phosphine at an abundance of 20 parts per billion (ppb). The authors checked all known abiotic methods for creating phosphine—photochemistry, volcanoes, meteoritic delivery—and determined that the phosphine originated either from an unknown abiotic pathway, or potentially a biotic one. News outlets quickly picked up the story, emphasising the possibility of life, with the Los Angeles Times running an article entitled "In the clouds of Venus, scientists may have found signs of extraterrestrial life" (Netburn, 2020). A month later, a different team led by Villanueva reanalysed the original data and published an article claiming Greaves et al. had committed calibration errors and their claimed phosphine result was actually caused by sulphur dioxide (Villanueva et al., 2021). This then led the original team to publish a re-analysis of their data maintaining that phosphine did exist in the Venusian cloud decks, albeit at lower levels than they had originally announced (Greaves et al., 2021). This response was published near-simultaneously with a paper by yet another team, led by Snellen, that analysed the data from one of Greaves et al.'s sources (ALMA, the Ata-

cama Large Millimeter/submillimeter Array) and found no statistically significant detection of phosphine (Snellen, I. A. G. et al., 2020). An abundance in the range of ppb is far smaller than could be detected in an exoplanetary atmosphere given current technology, but this example still illustrates the difficulty and ambiguity in the search for biosignatures, even in the solar system, cosmically-speaking our own backyard. I want to emphasise that no one on Greaves' team claimed to have discovered life on Venus, merely a candidate biosignature—and yet, the flurry of response papers shows the strong interest the subject of biosignatures evokes and the importance of rigour when working within the N=1 Problem.

The N=1 Problem is unavoidable in exoplanet science and astrobiology, and in representations of exoplanets, such as those found in science fiction. This is especially the case with SF featuring extra-terrestrial life. With only Earth life to extrapolate from, SF creators will often base their fictional extra-terrestrials on features of Earth life. Even the stranger fictional extra-terrestrials, such as the heptapods in *Arrival*, will still have recognisable features such as feet, or body parts viewers can liken to feet. There are also examples of science fictional works addressing the N=1 Problem itself. Asimov invokes the N=1 Problem in his novella *Nightfall*, about a planet in a system with six suns on which people only experience darkness once every thousand years. One character notes, discussing a planet that, like Earth, orbits only one sun "'Of course,' continued Beenay, 'there's the catch that life would be impossible on such a planet. It wouldn't get enough heat and light, and if it rotated there would be total Darkness half of each day. You couldn't expect life—which is fundamentally dependent upon light—to develop under those conditions" (Asimov, 1941). Because the characters in *Nightfall* are basing their definition of life on what they require, they consider the environment that the reader knows humans live in to be inhospitable. In this science fiction example, characters make an incorrect statement because they have only one example of life. Lem's *Solaris* also features (this time human) characters grappling with the N=1 Problem, as they encounter extra-terrestrial life for the first time and it takes them decades to confirm it is life. Lem's portrayal speculates that even with a second example of life, definitions may remain a challenge. This state of limbo is laid out: "The very existence of the thinking colossus would never let people abide in peace again. However much they travelled across the Galaxy and made contact with civilizations of other beings similar to us, Solaris would present a perpetual challenge to humankind" (Lem, 2014). The initial excitement, continuous reanalyses, and eventual consternation of the Solaricists occurs even with an entire alien planet and intelligence available to them. SF is

influenced by and provides commentary on the current state of science and science as a process—a challenge for science communication through conventional media—and this includes the N=1 Problem.

Similarly, it is difficult to avoid Earth (and by extension solar system) comparisons when discussing exoplanets. Messeri, the anthropologist studying exoplanet scientists referenced in Chapter 1, relays the efforts of astronomer Sara Seager and her team in deciding on terminology for an article on GJ1214b, an exoplanet belonging to that class intermediate between Earth and Neptune in mass and which has no analogue in our solar system. While Seager and her team repeatedly assert a desire to avoid Earth or solar-system-centric language due to the nature of GJ1214b, Messeri reports that "each time they tried to get away from solar system analogies, it became apparent that analogy was the only way out of the semantic gap" (Messeri, 2016). In searching for titles for possible compositions of GJ1214b, they ended up using solar system comparisons with "Mini-Neptune", "Water Planet", and "Super-Earth". Messeri writes that after a referee report suggested avoiding solar system analogies, these were changed to "Gas-Ice-Rock Planet with Primordial Gas Envelope", "Ice-Rock Planet with Sublimated Vapor Envelope", and "Rocky Planet with Outgassed Atmosphere". That these final titles are long and unwieldy speaks to exoplanet science's current lack of a standardized planetary taxonomy. This is an example of the N=1 Problem affecting discussions and naming conventions in exoplanet science.

Based on these examples from the definitions of life, biosignature research, and the science fiction literature, I argue that the N=1 Problem is a challenge for astrobiologists and for exoplanet scientists who work on research related to the search for life. The question then becomes, for exoplanet scientists: how does the N=1 Problem effect the search for habitable exoplanets, and the broader search for signs of life in the universe? Searching the literature shows two general methods researchers tend to follow that I will now introduce: either attempting to take non-Earth-centric approaches to life detection, or using features of Earth life to guide their work and search methodology. A "non-Earth-centric" approach means avoiding defining life and habitability by features that may be unique to Earth, and is difficult given the N=1 Problem. In exoplanet science, an example that might be considered Earth-centric would be looking for biosignatures only on exoplanets orbiting type G (Sun-like) stars at 1 AU (Earth-Sun distance)—this example is of course hypothetical as such planets have not yet been found. Already in the first issue of the journal *Astrobiology* in 2001, there were research

articles outlining ways to avoid Earth-centrism in astrobiology research, highlighting "The ultimate goal of a comprehensive life detection strategy is never to miss life when we encounter it" (Conrad and Nealson, 2001). Conrad and Nealson suggest that ideally scientists would search for generic manifestations of life's traits and requirements, like a solvent or limited reactivity of life's building blocks. Their paper demonstrates that this was an early priority in the field (Conrad and Nealson, 2001). The N=1 Problem makes it difficult to generalise about life, but there is a small technical literature about "weird life" or "life as we don't know it" (Schulze-Makuch and Irwin, 2008); (National Research Council, 2007); (Johnson et al., 2019), encouraged by late 20th century discoveries of extremophiles, or organisms that can thrive at extremes of temperature, salinity, pressure, radiation, etc. A well-known example of extremophiles is the bacteria living in hydrothermal vents on the ocean floor, where previously life had been unexpected due to the lack of light from the sun. Cleland suggests another way to avoid Earth-centrism specifically in assessing which traits of Earth life might be universal. She writes that since "microbes are the most representative form of life on Earth", having existed for billions of years through a wide range of atmospheric compositions and temperatures before comparatively fragile multicellular organisms with a narrower tolerable limit of parameters emerged, researchers who focus on multicellular life when discussing universal life are focusing on an unrepresentative sample even within our single example of life (Cleland, 2019). According to Cleland, researchers would be better served in assessing the universality of Earth life by focusing on microbes, as this would widen the range of environments life can tolerate.

The parameters of Earth life guide the search for Earth-like exoplanets, and for life in the universe. Examples include the focus on searching for planets in the habitable zone (Kasting et al., 1993). Working definitions of the habitable zone are based on habitable conditions for Earth life, and it is generally known as the distance from a star in which liquid water could be stable on a planetary surface with an Earth-like atmosphere. Searches for biosignatures also find target gases by considering what biosignatures Earth life might give off, such as oxygen and methane (Meadows et al., 2018). Schwieterman et al. base their list of candidate biosignature gases for exoplanet searches off of Earth's biosphere and the universality of the laws of chemistry, citing ozone, sulfur gases, and nitrous oxide as well as oxygen and methane as strong examples of biosignatures (Schwieterman et al., 2018). There is scientific evidence that indicates certain properties of Earth life might be universal. For example, the components

of water—oxygen and hydrogen—are readily available throughout the universe, and water is integral to life on Earth as the universal solvent (Ceccarelli, 2022), and has been found in exoplanet atmospheres (Tsiaras et al., 2019) and planet-forming disks (Hogerheijde et al., 2011). On Earth, carbon is essential to life because it has four valence electrons and can form strong bonds. Carbon is very common in the universe, having been found in asteroids and the interstellar medium, the dust between the stars (Henning and Salama, 1998), and potential 'carbon planets' have been discovered (Madhusudhan et al., 2012). Silicon also has four valence electrons and is very common in the universe and for this reason "silicon-based life" is a favourite of SF authors and has been speculated about by scientists as well (Cockell, 2016). However, silicon oxidises to a solid (becoming inert) and forms weaker bonds, therefore making carbon a better candidate for life chemically. Therefore, researchers like Cockell argue that carbon-based life with water as a universal solvent may well be universal (Cockell, 2016), and therefore provide a good basis for searches for life in the universe.

There are a lot of unknowns involved in the search for life in the universe, just as there are in related research into the origin of life. Researchers today are focusing on defining terms and assessment criteria, and determining *testable* hypotheses. Recently, Green et al. published a call for a framework for evidence for life beyond Earth, working to create a progressive scale to measure confidence of detections in the search for life in the universe (Green et al., 2021), and an earlier such scale was proposed by Almár and Race (Almár and Race, 2011). Rimmer investigates the potential for exoplanets to be laboratories for origin-of-life research, identifying "testability" as a central methodological challenge in origins work (Rimmer, 2023), and updating the Miller-Urey experiment for new environments. These are just a few examples of current and ongoing work to perform testable science in the search for life in the universe, a field which, because of the N=1 Problem, faces challenges in the realm of testability.

As much remains unknown—we simply do not have the data to address the N=1 Problem—philosophers also contribute heavily to discussions surrounding questions about the universal nature of life and the search for life in the universe. Indeed, it could be argued that some questions about life in the universe remain wholly in the realm of philosophy in the present. This has represented another challenge towards the respectability of the field of astrobiology as science. Billings investigated scientific and social perceptions of astrobiology as an established science, arguing that astrobiology has two narratives: a scientific one and a cultural one. As astrobiology has become an accepted multidisciplinary science, "The scientific story

of the search for life elsewhere is intricately intertwined with the story of how life began on Earth. This story is unfolding in the broader context of a story about who we are and where we are going, and why" (Billings, 2012). When scientific questions meet philosophical ones, as they do with the N=1 Problem, I argue it is even more important for scientists to be aware of the impact on their research. As Dick puts it "the extraterrestrial life debate in the twentieth century remains an example of science functioning at its limits" (Dick, 2020c).

To my knowledge, no social science study of the N=1 problem in exoplanet science of any scale has been attempted before. I therefore include this extensive background section to situate my interview results in the context of current landscape of exoplanet science and astrobiology. The N=1 Problem leads to a fundamentally biological question. Exoplanet scientists, though not biologists, have become involved in research related to it through searches for Earth-like exoplanets. The possibility of life in the universe has gone from entirely the realm of speculation to a topic of serious scientific study over recent decades, especially since the 1995 discovery of exoplanets. This discovery has resulted in substantial refocusing of scientific effort and funded missions looking for potential habitable exoplanets (Gillon, 2019). The focus on habitability involves exoplanet scientists in N=1: this research aims to investigate, through interviewing current exoplanet scientists, how N=1 affects their individual research and their field as a whole, looking to determine awareness and impact.

7.3 Interview Results

Here I present summations of the 18 participants' answers to each question in the interview, which asked after their awareness of the N=1 Problem and its impact on their research. The full list of questions asked, as well as the methodology for the interviews, is laid out in Chapter 3. The full interview transcripts are linked to in the Appendices. Due to the conversational framework of the interviews, results are collated and presented here question by question as a more readable form of my dataset. Researchers are pseudonymised by number, R1 to R18, and positions (academic levels), self-identified job titles, and areas of research are summarised in Table 7.1, which provides a common reference for each interviewee. Except for one, a geoscientist, all participants were exoplanet scientists or astronomers by self-identification.

Question: How does your work relate to the search for life in the universe?

This question places my sample of researchers into a broader context of exoplanet science

Table 7.1: Reference table for the researchers participating in the N=1 Problem study.

Researcher	Position	Job Title Self-ID
R1	Reader	Geochemist
R2	Associate Professor	Astronomer
R3	PhD Student	Astronomer
R4	PhD Student	Astronomer
R5	PhD Student	Astronomer
R6	Professor	Astronomer
R7	Postdoc	Astronomer, exoplanet scientist
R8	Postdoc	Astronomer
R9	Postdoc	Astronomer
R10	Reader	Astronomer
R11	Professor	Exoplanet scientist
R12	PhD Student	Astronomer, exoplanet scientist
R13	PhD Student	Astronomer, exoplanet scientist
R14	Lecturer	Astronomer
R15	PhD Student	Astronomer, exoplanet scientist
R16	Postdoc	Planetary astronomer
R17	Reader	Astronomer
R18	NASA scientist	Astronomer

by describing their research areas and contributions to the field.

Sample description of answers: The majority of participants stated they do not directly work on the search for life. A few participants work more directly on the processes that enable life—R2, a geoscientist, who works to replicate the early Earth environment and determine how life evolved; R9, who works on atmospheric evolution and planetary habitability; and R13, who works on lightning as a potential catalyst for life and the role of lightning in biosignature searches. Several are working on finding an Earth/Sun analogue system—R4 and R5, who are observationally looking for Earth twins, and R7, who works on the PLATO mission, describe such planets' potential to host life as the motivator for this work. However, the majority are not working directly on life detection but see their work as a related building block or first step towards life detection. For example, R6, who works on direct imaging of young Jupiter analogues orbiting close to their stars, said the techniques she uses will help us find Earth analogues, where there might be life. R12 works on hot Jupiters, considered unlikely to host life due to high temperatures, but says that the chemical processes she studies might also be applicable in determining what is and is not a sign of life. R16, who works on characterising the interiors of rocky planets, sees his work as a first step towards finding habitable conditions. R11, who works on numerical climate models of exoplanets, and R14, who works on planet formation, both stated their research did not originally relate to the search for life in the universe but this is an area they want to move towards. R8, who works on hot Jupiter atmospheres but is transitioning to rocky planets because she feels the field is moving in that direction, and R4 both said that finding life is the main goal of the field of exoplanet science. R10, who works on gravitational microlensing, expressed concerns that the field needs a new terminology for how we think about life.

What is your familiarity with the N=1 Problem?

None of the participants had heard the term "N=1 Problem", but all were familiar with the concept. Aside from R14, who said it came up in an undergraduate module while discussing the Drake equation, participants did not recall being formally taught the N=1 Probem during their education. Instead it's "always been there" (R8) or "obvious when you think about it" (R10). Participants often came across it in their research or in popular science books, like R7. Participants compared the N=1 Problem to the anthropic principle from cosmology (R6, R10, R14, R17), and the Drake equation (R14, R15), though recognised it was its own separate

concept. R11, R13, and R16 referred to it coming up most often in casual conversations with colleagues, at seminars, and at public outreach events.

How do you encounter the N=1 Problem in your work?

Several (R1, R3, R4, R9, R11, R13) said they encountered the N=1 Problem in their research in constraining requirements for life or habitable conditions, because they were basing these requirements on what is known of Earth life. R4 said the N=1 Problem was the motivation behind his research, rather than a challenge. R5 stated she doesn't yet encounter N=1 in her research, because that would be the next step after finding a habitable planet. R16 similarly said that his research is based on more defined laws of physics and chemistry, but that N=1 affects his larger research community. R8, R12, R15, and R17 also had work that was not directly related to life. R7, who works on PLATO, said that the N=1 Problem did not come up explicitly at work, but was the "unspoken driver". Similarly, R18 said the N=1 Problem comes up when designing fiducial planets for the HabWorlds observatory. R10 said he encounters the N=1 Problem in trying to fight off Earth-centric bias in his research, and R6 also expressed that she had had conversations about how taking the Earth as a model seemed dangerous.

Do you consider the N=1 Problem to be a problem (in the colloquial sense, for science)?

This question resulted in a wide variety of responses, with many participants considering ways in which N=1 both might and might not be seen as a problem. R2 felt N=1 limits the search because it already provides an answer, but acknowledged the challenge in how to identify life in the first place. R6 called it a prior, that is, a belief about the data: "In our inferences about life it's a very strong prior…if we had N=2 we still would have the same problem. I think about it in terms of detecting planets and how we would know if a planet actually is like the Earth. In the past few decades we've had several 'look, we've detected a habitable planet like the Earth!' but they're really not, they're mostly around lower-mass stars, very different conditions. I think in terms of finding habitable planets, finding planets like the earth, we won't just find one and be like, convincingly, 'this is life', we need to look at 30 or 40 to have a big enough sample to tell." R9 said N=1 was a problem in discussions rather than scientifically, because work is criticised for extrapolating from Earth life. He finds this extrapolation to not be a problem but rather the only way to perform work currently. R16 found N=1 problematic in how exoplanet researchers interface with policymakers and the public, who often end up under the impression this sort of research has already been done or

even that life has already been found, and this perception affects funding. R17 characterised N=1 as including the entire solar system (one example of a solar system with life), rather than just Earth, and said it was a problem because we have incomplete information about solar systems. All five of these are different reasons for calling N=1 a problem.

Similarly, several participants said N=1 was not a problem, also for a diverse set of reasons. R3 characterised it as an observation or a hindrance—to her, excitement for finding another Earth is not the inspiration for her interest in exoplanet science. R4 said N=1 is a "merit"; in giving us an example of a biosphere it gives us what to look for in searching for life in the universe. R8 said that she did not consider N=1 a problem or frustration, and did not think of it in a negative way. R11 said it was not a problem but a lack of understanding—and an unprovable hypothesis, in that to be certain about life in the universe every planet would have to be searched, which is impossible. R18 said that while extrapolations based on limited knowledge are a problem, overall N=1 is an exciting opportunity to learn something new.

Among the others, ideas brought up included that N=1 is only a problem in terms of lack of awareness of bias (R13), N=1 is a challenge (R10), N=1 will be a problem in the future but is not yet because of lack of data (R7), N=1 makes it hard to know how impactful research is due to the lack of overall data (R2), and N=1 might be a problem in that it effects mission design and might cause us to overlook interesting research areas, but there is no way around this so it is alright for now (R14).

Does the N=1 Problem help or hinder your research? Or both? Or neither?

This question was specified to be referring to the participant's research, not to their wider field. Help, hinder, both, and neither were all represented among participants' answers.

R1 felt N=1 helped by allowing more freedom and speculation in her research. R8 said N=1 helped by providing a starting point in the difficult task of proving one has found life. R10 and R15 said being aware of the N=1 Problem was helpful for their research. R17 said N=1 was useful to him to explain the impact of his research to others and explain why they should care about it.

For R2, N=1 hinders but is hard to get around. Relating to his answer to the previous question, R16 was concerned that public misunderstandings brought about by N=1 might negatively affect research funding.

There were common threads among those answering that N=1 both helps and hinders their research (R4, R7, R9, R11, R13, and R14): ways it helps include that Earth has a diversity of life and provides plenty to research already (R13), that N=1 provides a hook for funding applications (R7), that N=1 engages people to protect our one planet (R11), and that N=1 provides something to look for (R14). Ways N=1 hinders their research include a perception that most funding goes to Earth-centric exoplanet science projects (as opposed to, for example, hot Jupiter research) (R7), and several said that it is limiting and means scientists might overlook other forms of life.

Five participants (R3, R5, R6, R12, and R18) said that N=1 neither helps nor hinders their research, largely because their research is in areas less related to the detection of life. The prospect that it might affect their work in the future as more data is collected was brought up.

What would it mean for your work if life were found elsewhere in the solar system? What about on an exoplanet? Assume an unambiguous detection.

Positive feelings were generally ascribed to either scenario. "It would be amazing" said R2, but she was worried about trouble with societal implications and the strong potential for a disbelieving public. R5 said she would be excited either way but was not sure how it would affect her research, expressing a wish for detecting a wider range of exoplanets in general. R8 said it felt far away and was not something he had thought much about—either way, it would generate interest in solar system or exoplanet studies. R11 emphasised that at the moment we have not just N=1, but also "confirmed habitable conditions=1" so in that way we have a 100% success rate with finding life in habitable conditions. He said it might be more exciting to find a very different kind of life. R12 said that for either scenario, she would incorporate findings from that type of life into her models, for example atmospheric species observed on a life-bearing planet would become more important for her research. R16 said such a find would be game-changing, comparing it to exoplanet discovery in the 90s, and would result in Nobel prizes and increased funding. R18 said that either would be "a fundamental shift in our perception of our place in the universe", and said that, knowing astronomers, they would probably try and draw firm conclusions from this new sample size of 2—however, the second example would be sure to be different in some ways that would make people rethink things.

Regarding a find of life in the solar system, R4 said it would be difficult to prove this life had an independent origin from Earth life. R9 added that if such life were independent, it

would tell us a lot about how common life is and what environments it lives in. R8 felt a solar system discovery would be more exciting because it would be more "solid" and reachable. In the same vein, R10 brought up the possibility of a sample-return mission, though said this would not be near his line of research. Relatedly, R13 said it would not change his research much, because solar system life would likely be the sort that is not detectable from afar. R6 said a solar system find would inform the search for exoplanet life, for example, an increased focus on subsurface oceans if life were found on Europa. R15 said such a find would not affect his work unless the life were found in clouds, because then clouds would have to be considered as an environment for life. R14 and R17 stated a find would result in more funding for planetary science and research into lithopanspermia (rocks carrying life through space) and unusual environments, since it would prove life were possible on a very different environment than Earth.

It was commonly agreed a find of life on an exoplanet would direct more funding and scientific effort towards exoplanets. R9 said if life were found on an Earth-like exoplanet, it could indicate our type of life is common; if found in a different environment, we would have to rethink and expand our scientific methodology. R10 echoed that it would teach us a lot about where to find life. R3 would be especially interested if life were found on a sub-Neptune in the radius gap (a range of planetary radii slightly bigger than the Earth where there is evidence exoplanets are rare), as she focuses on such worlds. R15 said that for his specific research this would be N=1 two times rather than adding up to N=2, since "clouds are not really coupled to life on Earth". There was also skepticism an exoplanet find could be unambiguous. R4 said the challenge would be that it would take a very long time to confirm the detection. R6 did not think an unambiguous N=2 would be possible; there would be 10 ambiguous detections before one unambiguous. R13 and R14 mentioned it could be possible to unambiguously detect life if it was intelligent life and a technosignature, but ambiguous detections were more likely.

Do you believe that, with current technology, we can unambiguously detect life on another planet?

No participants believed that we could currently unambiguously detect life on another planet for certain. Some thought it was within the realm of possibility (R1, R2, R3, R5, R9, R15), and many gave emphatic "no"s (R4, R6, R7, R11, R12, R13, R16, R18). Several brought

up that unambiguously detecting intelligent life, i.e. a technosignature, would be far easier than confirming that a potential biosignature really did originate from life (R8, R10, R14, R17)—technosignatures could be certain, biosignatures must be ambiguous. Challenges include many possible contaminants and even though there are good candidate biosignatures, biosignatures do not equal life. Many answers considered the development of technology re: detecting biosignatures. JWST was referred to as a milestone (R3). R4 emphasised that we cannot even detect an exo-Earth yet (that is, an Earth-like planet in the habitable zone of a sun-like star), and we do not know how life started on earth, reflecting that "maybe we have a N=0.5 problem". Upcoming space missions that will add to our technical capability towards detecting biosignatures were cited—R6 mentioned LUVOIR (Large UV/Optical/IR Surveyor) and HabEx (Habitable Exoplanet Observatory), R13 added ARIEL (Atmospheric Remote-sensing Infrared Exoplanet Large-survey), and R18 added LIFE (Large Interferometer for Exoplanets) and R9 echoed this as those missions will have the ability to detect nitrogen in exoplanet atmospheres, currently beyond our capability. R9 mentioned the Venus phosphine findings to indicate the difficulty of biosignature detection even with technologically advanced instruments. R10 moved away from technology to say that the real problem was the lack of an accepted biological definition for life. R11 echoed that what is missing is the theory.

How do you manage the N=1 Problem in your research? Are you happy with how you handle it?

R3, R5, R6, R7, R12, R15, and R16 felt the N=1 Problem did not directly affect or impact their research, with R3 adding it "keeps things exciting".

Among those who did feel the N=1 Problem affected their research, there were a variety of responses. R4 said N=1 was the reason he did his research and did not feel it was something he had to "manage". Similarly, R9 was happy with his research, saying he "restricts himself to this environment [the Earth] because it's where we're experienced. People are often criticising but this is what we can certainly do and it doesn't exclude that other lifeforms exist." R11 said that he did not realise until this interview that N=1 guided his research, as he is basing his research on metabolisms on Earth metabolisms which he views as more solid research.

R1 found it helpful in addressing N=1 to "have a broad research base and spread out a bit". R2 considered herself open-minded as regards possibly habitable environments. R8 said that while N=1 provides a goal, she does not want to only search for things she expects, as

it is also interesting to investigate environments outside those expectations. R10 said that he drives his thinking separately from "prevalent astrobiology agendas" and is writing a new one of his own that focuses on emergence of life as a crucial factor to understand findings of future missions. R13 was happy with his awareness of N=1 in his research, saying his research group is careful to specify when talking about Earth life. R14 managed N=1 by collaborating with biologists and chemists, saying this was his solution for the fact that he felt unprepared by his physics and astronomy training for the biological N=1 problem. R17 used it as a motivation to study things outside Earth and the solar system, and was happy with this. R18 was happy with her day-to-day, but on her work on future missions she sees a need to solidly define the term "Earth-like".

Do you think astrobiologists and exoplanet scientists overall manage N=1 well? If not, how could they improve?

Generally, participants felt progress has been made on this, but also that there is more to be done. R1, R13, R14, and R16 felt interdisciplinary communication between astronomers and astrobiologists and geoscientists has improved in recent years, but could still be better, with R14 and R16 adding that grants and publications are often not set up for such interdisciplinary work and more venues for this should be established. R6 foresaw an even more interdisciplinary future for the two fields; stating the situation was fine for now because exoplanet science has not yet really had to grapple with N=1, but that in 20 years exoplanet science and astrobiology will merge. R12 and R17 praised astrobiology's extremophile research and its applications to space, with R12 adding that otherwise overall she felt we were limited in what we are looking for. R15 felt astrobiology did an excellent job combining the complexities of physics and biology, with biomarker research handled well because it does not discount other kinds of life while working within Earth-centric parameters due to scientific evidence supporting water and carbon. Concerns about N=1 were among the factors that inspired R10 to help found an interdisciplinary research centre.

R11 thought overall the community does "okay" but was worried about chasing sensational results and communicating uncertainty to the public. He noted that most of our communications implicitly assume N is greater than 1. R18 also felt scientists make strong statements due to public pressure, and sees what is important as making sure there is vigorous internal discussion and internal and external communications are well-balanced and different.

R2 felt N=1 is not given enough importance, and researchers are not open to serendipity. R3 felt some groups focus too much on our definition of life, instead there could be an "N=N" problem where every instance of life is so unique as to be not able to be replicated. R4 said some researchers do manage it well and some do not—he personally thinks, for example, looking for life around M-dwarfs is likely to be fruitless, saying "it's rightful and fundamentally scientific to be Earth-centred in our research because it's based on evidence". R5 said we should not be constrained to only looking for Earth life but should let Earth life shed light on our discoveries. R7 was happy and slightly frustrated—he mentioned reading the argument that "we are limiting ourselves to looking for ourselves" in a popular science book, but found that argument in itself limiting due jointly to the diversity of Earth's biosphere and the necessity of finding bounds for the problem. R17 noted that tech limitations will restrict astronomers at first to locations that are not Earth analogues, like red dwarf systems, saying "I think it's fine to shoot for one Earth mass at one AU and acknowledge that there will be other planets along the way".

R8 and R9 were not sure how to answer, expressing they would need to think about it more.

How does the N=1 Problem make you feel? What emotions does it bring up?

The N=1 Problem brought up a range of emotions in participants. Philosophical questions like "are we alone?" were commonly referenced, including by R1, R2, R3, R8, R12, R14, R16, and R18. Only R16, R15, R13, and R6 did not associate strong emotions with the N=1 Problem, viewing it from a more detached perspective. The most common emotion brought up was excitement—R2, R5, R7, R14, R17, and R18 brought up excitement specifically. R4 brought up a sense of pride at life itself. R17 was excited but also brought up a pessimism associated with thinking about how far we have to go to find answers relating to the N=1 Problem. R11 said he felt "interested, confused, emotionally uneasy and discombobulated", and R10 felt "puzzled, angry, and helpless" but also "inspired to think how we can overcome it."

7.4   Analysis

The interview questions were designed to determine the awareness and impact of the N=1 Problem on exoplanet science. While 100% of participants were familiar with the concept of

the N=1 Problem, only three out of the 18 scientists interviewed had heard of it by this name. None recalled having encountered it formally in their scientific education, except tangentially by related concepts like the Fermi Paradox, which asks why we have not yet found evidence of life in the universe, the Drake equation, which provides a framework for thinking about the prevalence of life in the universe, or the anthropic principle from cosmology. This lack of formal education, or even of having a name for the concept, suggests a role for the N=1 Problem in exoplanet science where it is prevalent but not explicitly discussed in academic contexts. Indeed, most participants referenced encountering it in popular science books, in fora where they interface with the public, or in less-formal academic settings like coffee breaks. This is in keeping with the consideration that N=1 might remain in the realm of philosophy for now, and with the assertion from several participants that the N=1 Problem is not something exoplanet science has encountered *yet*, for lack of data on habitable exoplanets. Among this sample of interviewed exoplanet science researchers, awareness of the N=1 Problem was universal.
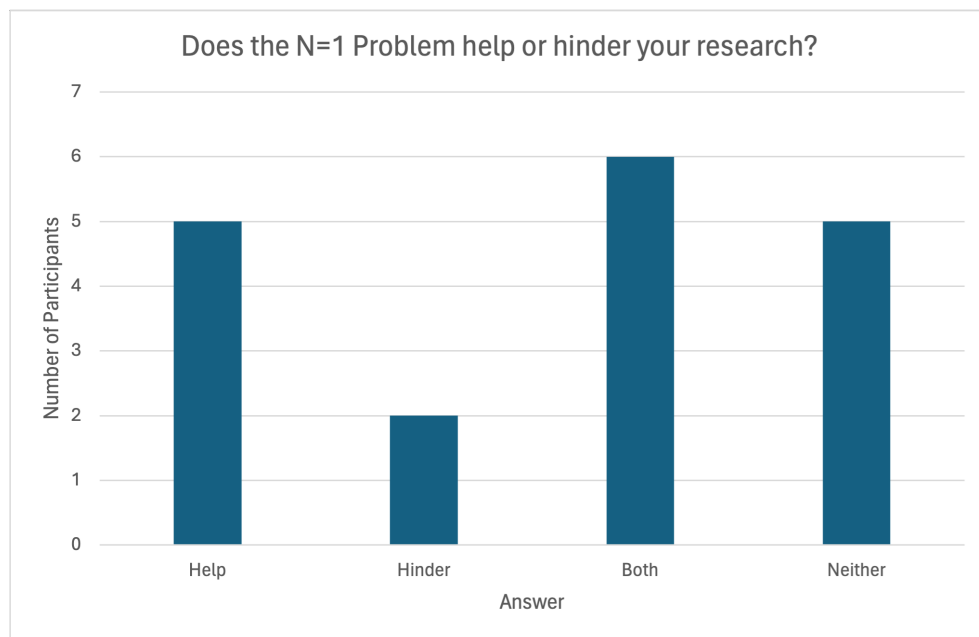
Figure 7.1: Participants' responses to the question "Does the N=1 Problem help or hinder your research?"

Findings related to impact were more varied. While participants thought the N=1 Problem was relevant to the field of exoplanet science, many stated their work was not directly related to life detection, with seven out of 18 participants saying the N=1 Problem was not something they directly dealt with in their research. Responses to the question "Does the N=1 Problem help or hinder your research?" were slightly different, as can be seen in Figure 7.1,

with answers spread over all four possible responses: help, hinder, both, or neither, and only five respondents answering "neither". Combining the "both" column with "help" and "hinder" in turn, it is apparent that overall more participants found the N=1 Problem to help their work. Figure 7.2 shows that even numbers of participants found the N=1 Problem to be a "problem" in the colloquial sense.
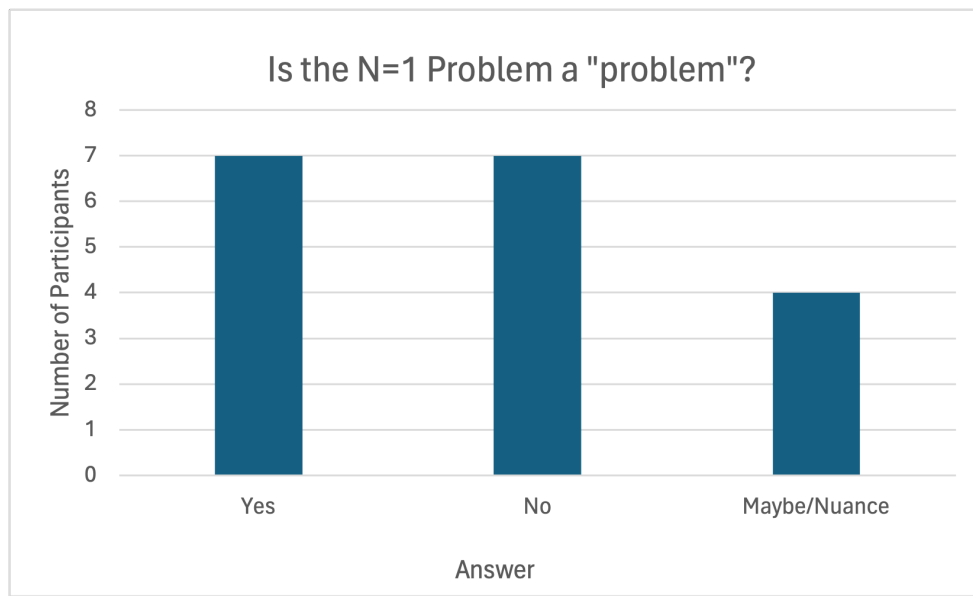
Figure 7.2: Participants' responses to the question "Is the N=1 Problem a 'problem'?" (in the colloquial sense).

Participants had a range of reactions to the N=1 Problem in their work. Not everyone felt it affected their research at all. Some thought it limited their research, and others thought it guided it. Two participants indicated during the interviews that the act of answering the questions made them realise that the N=1 Problem was guiding their choices of parameters used in their own methodologies. Through the interviews it was apparent many participants were grappling with not wanting to only search for Earth life out of concern they would limit themselves, but with having no other evidence-based research path upon which to embark. Several concluded that research based on Earth life was the way forward, but that an open mind must be maintained. Others experienced backlash for being "Earth-centric" when they felt their research was fundamentally scientific for being evidence-based. Still others looked for testable, scientific ways to avoid Earth-centrism in their research. It is clear that discussions are being had about this topic among exoplanet scientists, often in less-formal settings than their research papers. All had considered this problem before, and participants came up with several reframings of the N=1 Problem during the interviews:

- *The N=0.5 Problem*: refers to the fact we do not fully understand how Earth life started—if we don't understand our one example, we cannot have N=1.

- *The N=2 Problem*: finding just one more example of a living planet, and therefore having N=2, will not solve our sample size problem due to still being small number statistics.

- *The N=N Problem*: a challenge in which each example of life found is so different from all other examples that they cannot be compared.

These reframings indicate that participants in this study considered the complexity of the search for and definition of life, and were already thinking ahead to future challenges if the N=1 problem might be "solved" by finding another example of life.

Responses to the question *Do you believe that, with current technology, we can unambiguously detect life on another planet?*, shown in Figure 7.3, were universally doubtful or negative (except in the case of technosignatures), indicating that exoplanet scientists believe it will be a long time before the N=1 Problem might be "solved" by finding life on another planet. As discussed in the Background section, the N=1 Problem is often considered to fall into the realm of philosophy due to lack of empirical evidence, and considering it for their research makes astronomers think about the philosophical side of their work. Indeed, Billings' concept of the cultural aspect of research relating to life in the universe was invoked up by many participants, as they related that N=1 often came up in science communication, in speaking with the public, and in speaking with media and funding bodies—the search for life is part of the cultural story of their research.

Table 7.2 groups the positive and negative emotion words brought up when participants were asked whether the N=1 Problem brought up any feelings. Consistently, positive emotions associated with the N=1 Problem were related to excitement and potential—the N=1 Problem is connected to a great unknown, the question "Are we alone?", and that elicits a huge array of possibilities. The one outlier, "safe", referred to taking comfort in the idea that we might not find other life in the universe for a long time, so we are safe on Earth for now. "Cherishing" also referred to valuing our own planet. Negative emotions were largely ones of frustration and distress due to a lack of information about what life might be like in the universe. The "Other" category includes any word that is neither obviously positive nor obviously negative, and in this case is mostly composed of words describing the contemplative state of thought thinking about the N=1 Problem can induce. A commonality between all these emotion words is that
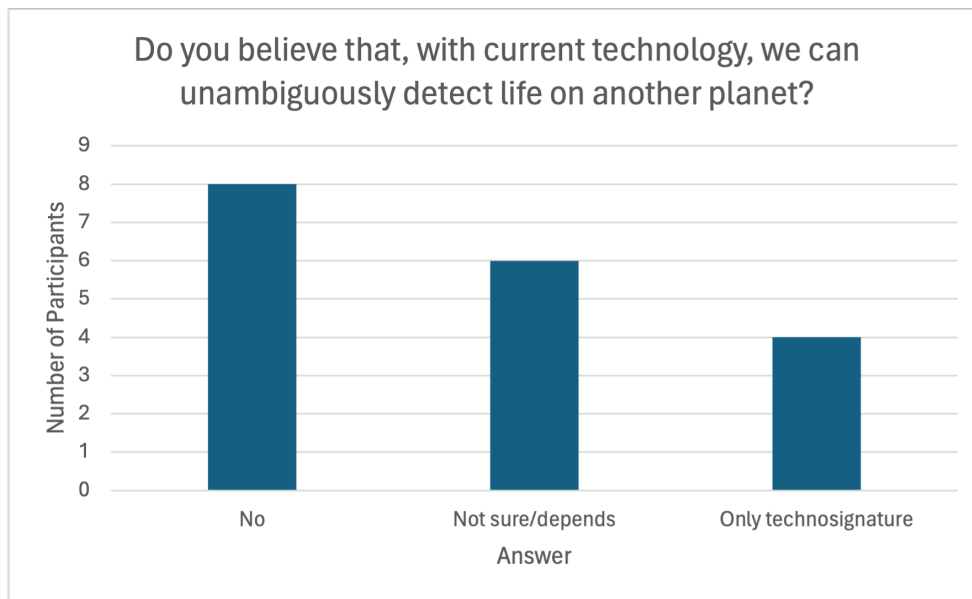
Figure 7.3: Participants' responses to the question "Do you believe that, with current technology, we can unambiguously detect life on another planet?".

they are all triggered by the *unknown* represented by the N=1 Problem. For the participants, that lack of knowledge can be exciting in its possibility, or overwhelming, or can make them ponder the big questions.

Participants were asked about how they managed the N=1 Problem in their own work. Among those participants who stated the N=1 Problem affects their research, several strategies were employed to manage it:

- Having a broad, flexible research base

- Staying open-minded

- Seeing finding a habitable planet as a motivating goal

- Learning more about our planet and life on Earth

- Being specific when talking about Earth life

- Rethinking astrobiology agendas

- Using N=1 to develop a clear-cut methodology (i.e. informing parameters) and basing research on life we know exists

- Collaborating with biologists and chemists

Table 7.2: Emotion words brought up in response to the N=1 Problem by participants.

Emotion words ascribed to the N=1 Problem

Positive	Negative	Other
Adventure	Angry	Contemplative
Anticipation	Caution	Philosophical
Cherishing	Concern	Reflective
Curious	Confused	Unemotional
Enthusiastic	Discombobulated	
Excited	Helpless	
Faith	Overwhelming	
Inspired	Puzzled	
Interested	Uneasy	
Safe	Overwhelming	
Wonder		

Not every participant would advocate every strategy; some felt using Earth-based parameters in exoplanet searches would unduly limit results. When asked how the field *overall* might improve management of the N=1 Problem, answers focused more on increasing interdisciplinary opportunities:

- Better dialogue between astronomers/astrobiologists/geoscientists

- Being open to serendipity

- Studying K-type stars as environments for life (stars intermediary in mass and temperature between M dwarfs and G type stars like the Sun)

- Not expecting to find life exactly like Earth life

- Increasing research into Earth's biosphere

- Increasing research into alternative chemistries

- Applying understandings from extremophile research to space

- Preparing to first observe biosignatures from planets around M dwarfs, due to technical limitations

- More funding and publishing schemes for interdisciplinary work

- More interdisciplinary research centres

- More conferences for interdisciplinary work

- Improving public communication

Largely, participants felt that most challenges arise in collaborations with other fields, and in communication to the public and to funders, rather than in the methodologies of current scientific endeavours. Exoplanet scientists found much to praise in current astrobiology research, and expressed desires for more interdisciplinary infrastructure to collaborate between their fields.

7.5   Conclusions

The goal of this study was to address the research question *In what ways does the N=1 Problem affect exoplanet science research?*. The N=1 Problem, a fundamentally biological problem, is a consideration in the field of exoplanet science through the search for habitable exoplanets. Interview results indicate that while there is a level of awareness of it among exoplanet scientists, it is not typically formally taught to them, nor has it received much attention in formal academic publishing related to exoplanet science. Interviewees see increased interdisciplinarity, especially with astrobiologists, as good practice moving forward in working within the N=1 Problem. This study sheds light on an aspect of exoplanet science not often discussed in academic research—as evidenced by participants not knowing the term for it, even if they were familiar with the concept—but that nevertheless impacts current exoplanet science research today, through setting parameters and influencing funding decisions. Concerns about a self-selection bias were mitigated when the participants ended up having a wide variety of views on the N=1 Problem (for example, participants who advocate for non-Earth-centrism in research and participants who use Earth parameters to search for exo-Earths were both represented). The goal of this study is not to evidence how common certain viewpoints are, as the sample size is not high enough, but rather to indicate a range of ideas present in the exoplanet science community. Results are therefore generalisable as a set of impacts N=1 *can* have on exoplanet scientists, but not as all impacts that N=1 has on exoplanet scientists.

8

Conclusions

This thesis has used quantitative and qualitative methodologies to study the portrayal of exoplanets in science fiction, finding a mutual influence between science and SF and suggesting a strong potential for the use of SF in teaching science. Science fiction is a genre of "what-ifs", which takes inspiration from science and extrapolates possible futures. It utilises cognitive estrangement to transport readers from their usual surroundings and explores the ramifications of science. Given the popularity of the genre and evidence supporting the pedagogical use of narrative, in order to determine SF's appropriateness for science education, I investigated the genre's portrayal of science. Exoplanet science was chosen because exoplanets have been portrayed in SF stories long before the empirical discovery of planets outside our solar system in the 1990s, providing a large corpus of science fiction exoplanets from before and after a scientific advance. I created a database of 212 science fiction exoplanets and used Bayesian network analysis to uncover trends in how various characteristics of SF exoplanets influence each other. The Bayesian network results show that there is a set of Earth-like characteristics authors frequently combine when creating SF exoplanets, and that in SF works from

after 1995, fictional exoplanets with intelligent native life or with established populations of non-native humans are less likely to appear. My database of fictional exoplanets are much more Earth-like than the population of real exoplanets, likely due to narrative demands of stories that largely feature human characters as well as due to the effect of the N=1 Problem. However, this trend towards less human-habitable planets among the post-1995 fictional exoplanets provides evidence that SF responds to scientific discoveries, as authors may have been exposed to reported scientific findings of exotic exoplanets like hot Jupiters and lava worlds, where humans could not survive. Bayesian network results provide statistical evidence that SF incorporates rapidly-evolving science, revealing trends in science fiction exoplanets that can guide further research and SF use in science education.

I also completed two interview research projects using qualitative methodology. During the creation of a science fiction anthology by the St Andrews Centre for Exoplanet Science (StACES), a research project was undertaken to understand the role of science in the decision-making process in SF worldbuilding. Stories were written by teams of scientists and writers, with the scientist acting as science consultant and playing a significant role in story generation. Qualitative survey data from project meetings were analysed, finding that scientific accuracy influenced plot decision-making and that participants saw the role of science in the stories as providing a jumping-off point and source of inspiration to readers to further explore scientific questions. The N=1 Project interviewed 18 scientists in the field of exoplanet science about the effects of the N=1 Problem, the fact that we only know of one planet with life, on their research. Results indicated that responses among participants varied, with some using N=1 to provide parameters in the search for life in the universe and others finding that N=1 limited research. Participants were frustrated by the unknowns of the N= Problem, and suggested increased interdisciplinary collaborations with other fields, particularly astrobiology, to help overcome the challenge. Both interview projects should be seen in light of the small sample size of interviewees and results should not be generalised beyond these circumstances. These two projects inform the portrayal of science in SF and its potential use for science education by evidencing the role of science in SF worldbuilding, and by emphasising that using Earth as an example happens in SF and science both—indeed, in most human endeavours of envisioning and understanding exoplanets.

This is an interdisciplinary thesis about the interplay between science, specifically exoplanet science, and science fiction. My research questions necessitated new methodologies and

I have created these from the disciplines at my disposal: astronomy, literature, and computational biology. Results are of interest to science communicators, particularly those involved in science education, as I used interdisciplinary methodologies to show that science and science fiction are not opposites, nor is theirs a one-way relationship: they mutually influence each other, in a sense they "model" for each other. Studying science fiction using a larger data set than the traditional close reading methodology allows for larger trends to become apparent, expanding the toolset for digital humanities work. Science communicators are already using innovative virtual tools and using the pedagogical power of storytelling in their work; it is my hope that the Bayesian network results can further these advances to include more science fiction especially in education. Next steps from my research could include writing lesson plans, investigating hard versus soft science fiction, or the role of terraforming on science fiction exoplanets, for example. Studying exoplanet science as it is portrayed in science fiction is valuable to astronomers because it provides evidence of the version of their science most accessible to the public—I argue scientists interested in public outreach and sharing their work with wider audiences would be benefited by following SF portrayals of their field. In response to C.P. Snow's thesis in *The Two Cultures* that there is a dangerous divide between the sciences and humanities in academia, my investigation of the mutual feedback between science and science fiction (humanities) suggests SF as one area in which the two cultures combine, or at least collaborate (Snow, 1959). My work therefore contributes to diverse fields of research. Ideally, this thesis will advocate for more interdisciplinary work of this nature.

A recurring theme through this thesis has been that the more unknowns there are in a field of science, the more likely researchers are to look to interdisciplinary collaboration to combine expertise to address the problem, and the more likely they are to reference the sort of applied speculation so common in SF. Exoplanet science is a field at the cutting-edge of our understanding of the universe, filled with unknowns (as shown by the N=1 Problem), and does indeed frequently collaborate with fields like geology, planetary science, atmospheric science, astrobiology, and more. Science fiction, as well, deals in the unknown. Novels purposely leave questions unanswered to create a narrative tension—this negative space fires the reader's mind to fill it in with possibilities. I argue that this element is a part of why science fiction is documented to inspire so many scientists into their fields to try and answer these unknowns. Science, of course, unlike fiction, aims to find answers by collecting evidence. More exploration of the effects of *the unknown* on both science and SF could be another fruitful avenue for future

research.

Science fiction is valuable not only as art, but also for science. It provides a human perspective on scientific advances, and sometimes an ethical check, as well as in some cases acting as the public's window into scientific progress and possibilities. Through this research, I have used Bayesian network and interview methodologies studying the portrayal of exoplanets in science fiction to show that changes in science affect SF, and that SF has strong potential for science education.

A

Data sets

Research data underpinning this thesis are available at:

- https://doi.org/10.17630/64ce9540-7c2b-4794-ace3-5d10b8f0d91c (data set of fictional exoplanets)

- https://doi.org/10.17630/e77b3d39-282c-4900-b89b-273a4c28654a (data set of anthology questionnaires)

- https://doi.org/10.17630/5982566e-8d7b-4a69-87ac-c40c31189539 (data set of N=1 Problem interview transcripts)

B

Settings file for Banjo

Input parameter settings file for

BA Bayesian

N Network Inference

J with Java

O Objects

Banjo is licensed from Duke University.

Settings file consistent with version 2.0.0

————————————————————-

———————————————————-

Project information

——————————————————-

project = fictional exoplanets

user = demo

dataset = ficexo22nov23.txt

notes = static bayesian network inference

——————————————————-

Search component specifications

——————————————————-

searcherChoice = Greedy

proposerChoice = RandomLocalMove

evaluatorChoice = default

newline deciderChoice = default

——————————————————-

Input and output locations

——————————————————-

inputDirectory = data

observationsFile = ficexo22nov23.txt

outputDirectory = output

reportFile = [insert file name here].txt

variableNames = InFile

——————————————————-

We require this only to validate the input

——————————————————-

———————————————————————————————

variableCount = 9

observationCount = 212

————————————————————————-

Pre-processing options

————————————————————————

discretizationPolicy =

discretizationExceptions =

createDiscretizationReport = standard

————————————————————————-

Network structure properties

————————————————————————-

minMarkovLag = 0

maxMarkovLag = 0

dbnMandatoryIdentityLags =

equivalentSampleSize = 1.0

maxParentCount = 2

defaultMaxParentCount = 10

————————————————————————-

Network structure properties, optional

————————————————————————-

variablesAreInRows = no

initialStructureFile =

mustBePresentEdgesFile =

mustNotBePresentEdgesFile =

————————————————————————-

Stopping criteria

————————————————-

maxTime = 15 m

maxProposedNetworks =

maxRestarts =

minNetworksBeforeChecking = 1000

————————————————-

Search monitoring properties

————————————————-

nBestNetworks = 10

bestNetworksAre = nonequivalent

screenReportingInterval = 20 s

fileReportingInterval = 5 m

————————————————-

Parameters used by specific search methods

————————————————-

For simulated annealing:

initialTemperature = 10000

coolingFactor = 0.7

reannealingTemperature = 800

maxAcceptedNetworksBeforeCooling = 2500

maxProposedNetworksBeforeCooling = 10000

minAcceptedNetworksBeforeReannealing = 500

For greedy:

minProposedNetworksAfterHighScore = 1000

minProposedNetworksBeforeRestart = 3000

maxProposedNetworksBeforeRestart = 5000

restartWithRandomNetwork = yes

maxParentCountForRestart = 2

————————————————————-

Command line user interface options

————————————————————-

askToVerifySettings = no

————————————————————-

Post-processing options

————————————————————-

createDotOutput = yes

computeInfluenceScores = yes

computeConsensusGraph = yes

createConsensusGraphAsHtml = yes

fileNameForTopGraph = [insert file name here]

fileNameForConsensusGraph = [insert file name here]

dotGraphicsFormat = jpg

dotFileExtension = txt

htmlFileExtension = html

fullPathToDotExecutable = /usr/local/Cellar/graphviz/2.44.1/bin/dot

timeStampFormat = yyyyMMddHHmmss

————————————————————-

Memory management and performance options

————————————————————-

precomputeLogGamma = yes

useCache = fastLevel2

cycleCheckingMethod = dfs

——————————————————————————-

Misc. options

——————————————————————————-
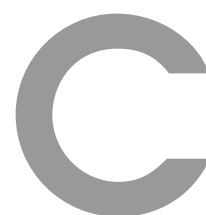
displayMemoryInfo = yes

displayStructures = yes

C

UTREC approval letter for StACES anthology project

School of Physics and Astronomy Ethics Committee

21 December 2020

Dear Christiane and Emma

Thank you for submitting your ethical application which was considered at the School Ethics Committee meeting on 21st Dec.

The School of Physics and Astronomy Ethics Committee, acting on behalf of the University Teaching and Research Ethics Committee (UTREC), has approved this application:

Approval Code:	PA15192	Approved on:	21st Dec 2020	Approval Expiry:	21st Dec 2025
Project Title:	St Andrews Centre for Exoplanet Science: Science Fiction Anthology				
Researcher(s):	Emma Puranen				
Supervisor(s):	Christiane Helling, Emily Finer, V Anne Smith				

The following supporting documents are also acknowledged and approved:

1. Application form
2. Participant consent
3. Participant information
4. Questionnaire
5. Sign up

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.
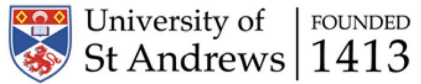
Yours sincerely,

Dimali Vithanage

School of Physics and Astronomy Ethics Committee
Dimali Vithanage, SEC administrator. Telephone: 01334 46 1682 Email: da16@st-andrews.ac.uk
The University of St Andrews is a charity registered in Scotland: No SC013532

Figure C.1: UTREC approval letter for StACES Anthology Project.

D

UTREC approval letter for N=1 project

School of Modern Languages Ethics Committee

14 March 2023

Dear Emma

Thank you for submitting your ethical application which has undergone proportionate review.

The School of Modern Languages Ethics Committee, acting on behalf of the University Teaching and Research Ethics Committee (UTREC), has approved this application:

Approval Code:	ML16820	Approved on:	13 March 2023	Approval Expiry:	13 March 2028
Project Title:	The N = 1 Problem in Exoplanet Science				
Researcher(s):	Emma Puranen				
Supervisor(s):	Dr Emily Finer				

The following supporting documents are also acknowledged and approved:

1. Participant Information Sheet
2. Consent Form
3. Sample Interview Questions

Approval is awarded for 5 years, see the approval expiry date 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 you must submit an ethical review application form.

If you need to make changes to your project, 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,

School of Modern Languages Ethics Committee
Dr Emily Finer/Leanne Bell, Buchanan Building, Union Street, St Andrews, Fife, KY16 9PH
Telephone: 01334 462949 Email: langsethics@st-andrews.ac.uk
The University of St Andrews is a charity registered in Scotland: No SC013532

Figure D.1: UTREC approval letter for N=1 project.

Bibliography

Abraham, D. (Writer). and Woolnough, J. (Director). (2021). The Expanse.

Adúriz-Bravo, A. (2014). Teaching the nature of science with scientific narratives. *Interchange*, 45.

Akiyama, K., Alberdi, A., Alef, W., Asada, K., Azulay, R., Baczko, A.-K., Ball, D., Baloković, M., Barrett, J., Bintley, D., Blackburn, L., Boland, W., Bouman, K. L., Bower, G. C., Bremer, M., Brinkerink, C. D., Brissenden, R., Britzen, S., Broderick, A. E., Broguiere, D., Bronzwaer, T., Byun, D.-Y., Carlstrom, J. E., Chael, A., kwan Chan, C., Chatterjee, S., Chatterjee, K., Chen, M.-T., Chen, Y., Cho, I., Christian, P., Conway, J. E., Cordes, J. M., Crew, G. B., Cui, Y., Davelaar, J., Laurentis, M. D., Deane, R., Dempsey, J., Desvignes, G., Dexter, J., Doeleman, S. S., Eatough, R. P., Falcke, H., Fish, V. L., Fomalont, E., Fraga-Encinas, R., Freeman, W. T., Friberg, P., Fromm, C. M., Gómez, J. L., Galison, P., Gammie, C. F., García, R., Gentaz, O., Georgiev, B., Goddi, C., Gold, R., Gu, M., Gurwell, M., Hada, K., Hecht, M. H., Hesper, R., Ho, L. C., Ho, P., Honma, M., Huang, C.-W. L., Huang, L., Hughes, D. H., Ikeda, S., Inoue, M., Issaoun, S., James, D. J., Jannuzi, B. T., Janssen, M., Jeter, B., Jiang, W., Johnson, M. D., Jorstad, S., Jung, T., Karami, M., Karuppusamy, R., Kawashima, T., Keating, G. K., Kettenis, M., Kim, J.-Y., Kim, J., Kim, J., Kino, M., Koay, J. Y., Koch, P. M., Koyama, S., Kramer, M., Kramer, C., Krichbaum, T. P., Kuo, C.-Y., Lauer, T. R., Lee, S.-S., Li, Y.-R., Li, Z., Lindqvist, M., Liu, K., Liuzzo, E., Lo, W.-P., Lobanov, A. P., Loinard, L., Lonsdale, C., Lu, R.-S., MacDonald, N. R., Mao, J., Markoff, S., Marrone, D. P., Marscher, A. P., Martí-Vidal, I., Matsushita, S., Matthews, L. D., Medeiros, L., Menten, K. M., Mizuno, Y., Mizuno, I., Moran, J. M., Moriyama, K., Moscibrodzka, M., Müller, C., Nagai, H., Nagar, N. M., Nakamura, M., Narayan, R., Narayanan, G., Natarajan, I., Neri, R., Ni, C., Noutsos, A., Okino, H., Olivares, H., Ortiz-León, G. N., Oyama, T., Özel, F., Palumbo, D. C. M., Patel, N., Pen, U.-L., Pesce, D. W., Piétu, V., Plambeck, R., PopStefanija, A., Porth, O., Prather, B., Preciado-López, J. A., Psaltis, D., Pu, H.-Y., Ramakrishnan, V., Rao, R., Rawlings, M. G., Raymond, A. W., Rezzolla, L., Ripperda, B., Roelofs, F., Rogers, A., Ros, E., Rose, M., Roshanineshat, A., Rottmann, H., Roy, A. L., Ruszczyk, C., Ryan, B. R., Rygl, K. L. J., Sánchez, S., Sánchez-Arguelles, D., Sasada, M., Savolainen, T., Schloerb, F. P., Schuster, K.-F., Shao, L., Shen, Z., Small, D., Sohn, B. W., SooHoo, J., Tazaki, F., Tiede, P., Tilanus, R. P. J., Titus, M., Toma, K., Torne, P., Trent, T., Trippe, S., Tsuda, S., van Bemmel, I., van Langevelde, H. J., van Rossum, D. R., Wagner, J., Wardle, J., Weintroub, J., Wex, N., Wharton, R., Wielgus, M., Wong, G. N., Wu, Q., Young, K., Young, A., Younsi, Z., Yuan, F., Yuan, Y.-F., Zensus, J. A., Zhao, G., Zhao, S.-S., Zhu, Z., Algaba, J.-C., Allardi, A., Amestica, R., Anczarski, J., Bach, U., Baganoff, F. K., Beaudoin, C., Benson, B. A., Berthold,

R., Blanchard, J. M., Blundell, R., Bustamente, S., Cappallo, R., Castillo-Domínguez, E., Chang, C.-C., Chang, S.-H., Chang, S.-C., Chen, C.-C., Chilson, R., Chuter, T. C., Rosado, R. C., Coulson, I. M., Crawford, T. M., Crowley, J., David, J., Derome, M., Dexter, M., Dornbusch, S., Dudevoir, K. A., Dzib, S. A., Eckart, A., Eckert, C., Erickson, N. R., Everett, W. B., Faber, A., Farah, J. R., Fath, V., Folkers, T. W., Forbes, D. C., Freund, R., Gómez-Ruiz, A. I., Gale, D. M., Gao, F., Geertsema, G., Graham, D. A., Greer, C. H., Grosslein, R., Gueth, F., Haggard, D., Halverson, N. W., Han, C.-C., Han, K.-C., Hao, J., Hasegawa, Y., Henning, J. W., Hernández-Gómez, A., Herrero-Illana, R., Heyminck, S., Hirota, A., Hoge, J., Huang, Y.-D., Impellizzeri, C. M. V., Jiang, H., Kamble, A., Keisler, R., Kimura, K., Kono, Y., Kubo, D., Kuroda, J., Lacasse, R., Laing, R. A., Leitch, E. M., Li, C.-T., Lin, L. C.-C., Liu, C.-T., Liu, K.-Y., Lu, L.-M., Marson, R. G., Martin-Cocher, P. L., Massingill, K. D., Matulonis, C., McColl, M. P., McWhirter, S. R., Messias, H., Meyer-Zhao, Z., Michalik, D., Montaña, A., Montgomerie, W., Mora-Klein, M., Muders, D., Nadolski, A., Navarro, S., Neilsen, J., Nguyen, C. H., Nishioka, H., Norton, T., Nowak, M. A., Nystrom, G., Ogawa, H., Oshiro, P., Oyama, T., Parsons, H., Paine, S. N., Peñalver, J., Phillips, N. M., Poirier, M., Pradel, N., Primiani, R. A., Raffin, P. A., Rahlin, A. S., Reiland, G., Risacher, C., Ruiz, I., Sáez-Madaín, A. F., Sassella, R., Schellart, P., Shaw, P., Silva, K. M., Shiokawa, H., Smith, D. R., Snow, W., Souccar, K., Sousa, D., Sridharan, T. K., Srinivasan, R., Stahm, W., Stark, A. A., Story, K., Timmer, S. T., Vertatschitsch, L., Walther, C., Wei, T.-S., Whitehorn, N., Whitney, A. R., Woody, D. P., Wouterloot, J. G. A., Wright, M., Yamaguchi, P., Yu, C.-Y., Zeballos, M., Zhang, S., and Ziurys, L. (2019). First M87 Event Horizon Telescope results. i. the shadow of the supermassive black hole. *The Astrophysical Journal Letters*, 875(1):L1.

Aldiss, B. (1973). *Billion year spree: the true history of science fiction*. Doubleday.

Almár, I. and Race, M. (2011). Discovery of extra-terrestrial life: assessment by scales of its importance and associated risks. *Philosophical Transactions of the Royal Society of Astronomy*, 369:679–692.

Alter, A. (2017). Andy Weir's best seller 'The Martian' gets a classroom-friendly makeover. *The New York Times*.

Anders, C. (2019a). The bizarre planets that could be humanity's new homes. *The Atlantic*.

Anders, C. (2019b). *The City in the Middle of the Night*. Tor Books.

Asimov, I. (1941). Nightfall. *Astounding Science Fiction*.

Asimov, I. (1975). How easy to see the future! *Natural History*.

Baldassano, C., Chen, J., Zadbood, A., Pillow, J., Hasson, U., and Norman, K. (2017). Discovering event structure in continuous narrative perception and memory. *Neuron*, 95(3).

Barnes, R. (2017). Tidal locking of habitable exoplanets. *Celestial Mechanics and Dynamical Astronomy*, 129(4):509–536.

Barnes, R., Jackson, B., Greenberg, R., and Raymond, S. N. (2009). Tidal limits to planetary habitability. *The Astrophysical Journal*, 700(1):L30–L33.

Baxter, S. (2014). Big planets: Super-Earths in science ficiton. *Journal of the British Interplanetary Society*, 67:105–109.

Beichman, C. A. and Green, T. P. (2018). Observing exoplanets with the James Webb Space Telescope. In Deeg, H. and Belmonte, J., editors, *Handbook of exoplanets*, pages 1283–1308. Springer.

Berlina, A. (2017). *Viktor Shklovsky: a reader*. Bloomsbury Academic.

Berns, G., Blaine, K., Prietula, M., and Pye, B. (2013). Short- and long-term effects of a novel on connectivity in the brain. *Brain Connectivity*, 3(6).

Billings, L. (2012). Astrobiology in culture: The search for extraterrestrial life as "science". *Astrobiology*, 12(10):966–75.

Bitsch, Bertram, Johansen, Anders, Lambrechts, Michiel, and Morbidelli, Alessandro (2015). The structure of protoplanetary discs around evolving young stars. *Astronomy & Astrophysics*, 575:A28.

Bonfanti, A., Fossati, L., Kubyshkina, D., and Cubillos, P. E. (2021). Constraining stellar rotation and planetary atmospheric evolution of a dozen systems hosting sub-neptunes and super-earths. *Astronomy & Astrophysics*, 656:A157.

Borucki, W. J., Koch, D., Basri, G., Batalha, N., Brown, T., Caldwell, D., Caldwell, J., and Christensen-Dalsgaard, J. (2010). Kepler planet-detection mission: introduction and first results. *Science*, 327(5968).

Bradley, J. (1727). A letter from the reverend mr james bradley savilian professor of astronomy at oxford, and f.r.s. to dr. edmond halley astronom. reg. &c. giving an account of a new discovered motion of the fix'd stars. *Philosophical Transactions of the Royal Society of London*.

Brennan, P. (2022). Cosmic milestone: NASA confirms 5,000 exoplanets. *NASA-JPL*.

Brennan, P. (2023). Life on other planets: what is life and what does it need? *NASA Exoplanet Exploration*.

Burton, R., editor (1885). *The Book of the Thousand Nights and a Night — Volume 05*. The Burton Club.

Calapez, F., Dias, R., Cesário, R., Gonçalves, D., Pedras, B., Canário, J., and Martins, Z. (2023). Spectroscopic detection of biosignatures in natural ice samples as a proxy for icy moons. *Life*, 13(2).

Campbell, B., Walker, G., and Yang, S. (1988). A Search for Substellar Companions to Solar-type Stars. *Astrophysical Journal*, 331.

Carroll, R. (2013). Kip Thorne: physicist studying time travel tapped for Hollywood film. *The Guardian*.

Carter, J. (2022). Welcome to Kepler-16b, a 'Tatooine-like' planet that orbits two suns newly spotted from Earth. *Forbes*.

Cavendish, M. (1668). *The Description of a New World, Called the Blazing-World*. A. Maxwell.

Ceccarelli, C. (2022). Water in the universe. In Gargaud, M., editor, *Encyclopedia of Astrobiology*. Springer.

Charbonneau, D., Brown, T. M., Latham, D. W., and Mayor, M. (1999). Detection of planetary transits across a sun-like star. *The Astrophysical Journal*, 529(1).

Chauvin, G., Lagrange, A.-M., Dumas, C., Zuckerman, B., Mouillet, D., Song, I., Beuzit, J.-L., and Lowrance, P. (2004). A giant planet candidate near a young brown dwarf. *Astronomy and Astrophysics*, 425(2):L29–L32.

Chickering, D. M. (1996). Learning Bayesian networks is NP-complete. In Fisher, D. and Lenz, H. J., editors, *Learning from data. Lecture notes in statistics, vol 112*, pages 121–130. Springer, New York, NY.

Childers, G., Governer, D., Greer, K., and James, V. (2023). Science fiction conventions as places to communicate science. *Research in Science Education*, 54:149–166.

Choi, C. Q. (2015). Scientists spot distant alien planet with a comet-like tail. *NBC News*.

Clarke, B. (2015). The planetary imaginary: Gaian ecologies from Dune to Neuromancer. In Clarke, B., editor, *Earth, Life, and System*, pages 151–174. Fordham University Press, New York, USA.

Cleland, C. E. (2019). Challenges for a universal theory of life. In Cleland, C., editor, *The Quest for a Universal Theory of Life: Searching for Life As We Don't Know It*, page 105–131. Cambridge University Press.

Cockell, C. (2016). The similarity of life across the universe. *Molecular Biology of the Cell*, 27:1553–1555.

Coffey, Y., Bhullar, N., Durkin, J., Islam, M. S., and Usher, K. (2021). Understanding eco-anxiety: a systematic scoping review of current literature and identified knowledge gaps. *The Journal of Climate Change and Health, 3*, 3(100047).

Coll Adanay, M. and Sporleder, C. (2015). Clustering of novels represented as social networks. *Linguistic Issues in Language Technology*, 12.

Conrad, P. and Nealson, K. (2001). A non-earthcentric approach to life detection. *Astrobiology*, 1:15–24.

Copernicus, N. (1543). *De Revolutionibus Orbium Cœlestium, Libri I-VI*. Octavo.

Cortesão, M. (2015). What is life: Various definitions towards the contemporary astrobiology.

Dawson, R. I. and Johnson, J. A. (2018). Origins of Hot Jupiters. *Annual Review of Astronomy and Astrophysics*, 56(1):175–221.

Dick, S. J. (2020a). Critical issues in the history, philosophy, and sociology of astrobiology. In Dick, S. J., editor, *Space, time, and aliens: collected works on cosmos and culture*, pages 655–684. Springer.

Dick, S. J. (2020b). The philosophy of astronomy, cosmology, and astrobiology: A preliminary reconnaissance. In Dick, S. J., editor, *Space, time, and aliens: collected works on cosmos and culture*, pages 631–650. Springer.

Dick, S. J. (2020c). Plurality of worlds: a persistent theme in western civilization. In Dick, S. J., editor, *Space, time, and aliens: collected works on cosmos and culture*, pages 5–25. Springer.

Dunér, D., Capova, K. A., Gargaud, M., Geppert, W., Kereszturi, A., and Persson, E. (2018). Astrobiology and society in europe. In Capova, K. A., Persson, E., Milligan, T., and Dunér, D., editors, *Astrobiology and Society in Europe Today*, pages 7–10. Springer International Publishing.

Eichmeier, A. A., Bao, L., Xenos, M. A., Brossard, D., and Scheufele, D. A. (2023). Fictional scenarios, real concerns: science fiction and perceptions of human genome editing. *JCOM*, 22.

Epicurus (1925). Letter to herodotus.

ESA Voyage 2050 (2021). Voyage 2050 sets sail: ESA chooses future science mission themes. *ESA Director's Desk*.

exoplanet.eu (2024). The extrasolar planets encyclopedia.

Finer, E. (2010). *Turning into Sterne: Viktor Shklovskii and literary reception*. MHRA Legenda.

Fischer, D. A., Howard, A. W., Laughlin, G. P., Macintosh, B., Mahadevan, S., Sahlmann, J., and Yee, J. C. (2014). Exoplanet detection techniques. In Beuther, H., editor, *Protostars and planets VI*, pages 715–737. University of Arizona.

Flood, A. (2019). 'Sci-fi makes you stupid' study refuted by scientists behind original research. *The Guardian*.

Freedman, C. (2000). *Critical Theory and Science Fiction*. Wesleyan University Press.

Friedman, N. (1998). The Bayesian structural EM algorithm. In *Proceedings of the fourteenth Conference on Uncertainty in Artificial Intelligence*, pages 129–138, San Francisco, CA, USA. Morgan Kaufmann Publishers Inc.

Fujii, Y., Angerhausen, D., Deitrick, R., Domagal-Goldman, S., Grenfell, J. L., Hori, Y., Kane, S. R., Palle, E., Rauer, H., Siegler, N., Stapelfeldt, K., and Stevenson, K. B. (2018). Exoplanet biosignatures: Observational prospects. *Astrobiology*, 18(6):739–778.

Gansner, E. R. and North, S. C. (2000). An open graph visualization system and its applications to software engineering. *Journal of Software: Practice and Experience*, 30(11).

Gavaler, C. and Johnson, D. (2017). The genre effect: A science fiction (vs. realism) manipulation decreases inference effort, reading comprehension, and perceptions of literary merit. *Scientific Study of Literature*, 7:79–108.

Gavaler, C. and Johnson, D. (2019). The literary genre effect : A one-word science fiction (vs. realism) manipulation reveals intrinsic text properties outweigh extrinsic expectations of literary quality. *Scientific Study of Literature*, 9(1):34–52.

Gillon, M. (2019). Life under another sun: From science fiction to science. *European Review*, 28(1):18–39.

Gillon, M., Triaud, A., Demory, B., Dehin, E., Agol, E., Deck, K., Lederer, S., de Wit, J., Burdanov, A., Ingalls, J., Bolmont, E., Leconte, J., Raymond, S., Selsis, F., Turbet, M., Barkaoui, K., Burgasser, A., Burleigh, M., Carey, S., Chaushev, A., Copperwheat, C., Delrez, L., Fernandes, C., Holdsworth, D., Kotze, E., Van Grootel, V., Almleaky, Y., Benkhaldoun, Z., Magain, P., and Queloz, D. (2017). Seven temperate terrestrial planets around the nearby ultracool dwarf star trappist-1. *Nature*, 542:456–460.

Good, O. (2017). Elite: Dangerous 'predicted' the discovery of the Trappist-1 system. *Polygon*.

Greaves, J., Richards, A., Bains, W., Rimmer, P., Clements, D., Seager, S., Pętkowski, J., Sousa-Silva, C., Ranjan, S., and Fraser, H. (2021). Reply to: No evidence of phosphine in the atmosphere of Venus from independent analyses. *Nature Astronomy*, 5.

Greaves, J., Richards, A., Bains, W., Rimmer, P., Sagawa, H., Clements, D., Seager, S., Petkowski, J., Sousa-Silva, C., Ranjan, S., Drabek-Maunder, E., Fraser, H., Cartwright, A., Mueller-Wodarg, I., Zhan, Z., Friberg, P., Coulson, I., Lee, E., and Hoge, J. (2021). Phosphine gas in the cloud decks of Venus. *Nature Astronomy*, 5:655–664.

Green, J. (2023). *The Possibility of Life: Science, Imagination, and Our Quest for Kinship in the Cosmos*. Hanover Square Press.

Green, J., Hoehler, T., Neveu, M., Domagal-Goldman, S., Scalice, D., and Voytek, M. (2021). Call for a framework for reporting evidence for life beyond Earth. *Nature*, pages 575–579.

Hartemink, A. J. (2010). Banjo: Bayesian network inference with Java objects version (2.2.0).

Heckerman, D., Geiger, D., and Chickering, D. M. (1995). Learning Bayesian networks: the combination of knowledge and statistical data. *Machine Learning*, 20:197–243.

Heller, R. and Armstrong, J. (2014). Superhabitable worlds. *Astrobiology*, 14(1).

Henning, T. and Salama, F. (1998). Carbon in the universe. *Science*, 282(5397):2204–2210.

Hogerheijde, M. R., Bergin, E. A., Brinch, C., Cleeves, L. I., Fogel, J. K. J., Blake, G. A., Dominik, C., Lis, D. C., Melnick, G., Neufeld, D., Panić, O., Pearson, J. C., Kristensen, L., Yıldız, U. A., and van Dishoeck, E. F. (2011). Detection of the water reservoir in a forming planetary system. *Science*, 334(6054):338–340.

Holmes, R. (2010). The Royal Society's lost women scientists. *The Guardian*.

Howell, S. B. (2020). The grand challenges of exoplanets. *Frontiers in Astronomy and Space Sciences*, 7.

Hsu, D. C., Ford, E. B., Ragozzine, D., and Ashby, K. (2019). Occurrence rates of planets orbiting FGK stars: Combining Kepler DR25, Gaia DR2, and Bayesian inference. *The Astronomical Journal*, 158(3).

Hui, E., Stafford, R., Matthews, I., and Smith, V. (2021). Bayesian networks as a novel tool to enhance interpretability and predictive power of ecological models. *Ecological Informatics*, 68.

Huygens, C. (1698). *Cosmotheoros*. Timothy Childe.

Irani, M. and Weitkamp, E. (2023). Factors affecting the efficacy of short stories as science communication tools. *JCOM*, 22(2).

James, O., von Tunzelmann, E., Franklin, P., and Thorne, K. (2015). Gravitational lensing by spinning black holes in astrophysics, and in the movie Interstellar. *Classical and Quantum Gravity*, 32.

Jemisin, N. (2018). Hugo acceptance speech.

Johnson, S. A., Penny, M., Gaudi, B. S., Kerins, E., Rattenbury, N. J., Robin, A. C., Novati, S. C., and Henderson, C. B. (2020). Predictions of the nancy grace roman space telescope galactic exoplanet survey. ii. free-floating planet detection rates*. *The Astronomical Journal*, 160.

Johnson, S. S., Graham, H., Anslyn, E., Conrad, P., Cronin, L., Ellington, A., Elsila, J., Girguis, P., House, C., Kempes, C., Libby, E., Mahaffy, P., Nadeau, J., Sherwood Lollar, B., Steele, A., Chou, L., Grefenstette, N., and Da Poian, V. (2019). Agnostic Approaches to Extant Life Detection. In LPI Editorial Board, editor, *Mars Extant Life: What's Next?*, volume 2108 of *LPI Contributions*, page 5026.

Jungnickel, D. (1999). The greedy algorithm. In Cook, W., editor, *Graphs, networks and algorithms: algorithms and computation in mathematics, volume 5*, pages 129–153. Springer.

Kaltenegger, L. (2017). How to characterize habitable worlds and signs of life. *Annual Review of Astronomy and Astrophysics*, 55(1):433–485.

Kane, S. (2023).

Kasting, J., Whitmire, D., and Reynolds, R. (1993). Habitable zones around main sequence stars. *Icarus*, 101(1):108–128.

Kepler, J. (1634). *Somnium*.

Khatchadourian, R. (2020). N.K. Jemisin's dream worlds. *The New Yorker*.

Kirkpatrick, S., Gelatt Jr, C., and Vecchi, M. (1983). Optimization by simulated annealing. *Science*, 220.

Kopparapu, R. K., Hébrard, E., Belikov, R., Batalha, N. M., Mulders, G. D., Stark, C., Teal, D., Domagal-Goldman, S., and Mandell, A. (2018). Exoplanet classification and yield estimates for direct imaging missions. *The Astrophysical Journal*, 856(2).

Kopparapu, R. K., Ramirez, R., Kasting, J. F., Eymet, V., Robinson, T. D., Mahadevan, S., Terrien, R. C., Domagal-Goldman, S., Meadows, V., and Deshpande, R. (2013). Habitable zones around main-sequence stars: new estimates. *The Astrophysical Journal*, 765(2).

Kowal, M. (2020). World-building: How science sculpts science fiction.

Krupp, E. (2015). Native American cosmology and other worlds. In Gargaud, M., editor, *Encyclopedia of Astrobiology*. Springer.

Lambie, R. (2017). Into the unknown: curator Patrick Gyger on his new sci-fi exhibition. *Den of Geek*.

Landrum, R., Brakke, K., and McCarthy, M. (2019). The pedagogical power of storytelling. *Scholarship of Teaching and Learning in Psychology*, 5(3):247–253.

Lapirov-Skoblo, M. (1924). Science and technology: Journeys in interplanetary space. *Pravda*.

Latham, D., Mazeh, T., Stefanik, R., Mayor, M., and Burki, G. (1989). The unseen companion of HD114762: a probable brown dwarf. *Nature*, 339:38–40.

Latour, B. (1999). *Pandora's Hope*. Harvard University Press.

Le Guin, U. (1969). *The Left Hand of Darkness*. Orion Publishing Group Ltd.

Lem, S. (2014). *Solaris*. Pro Auctore Wojciech Zemek.

Lingam, M. (2021). A brief history of the term 'habitable zone' in the 19th century. *International Journal of Astrobiology*, 20(5):332–336.

Lovis, C. and Fischer, D. (2010). Radial Velocity Techniques for Exoplanets. In Seager, S., editor, *Exoplanets*, pages 27–53. University of Arizona Press.

Lowe, T., Brown, K., Dessai, S., de França Doria, M., Haynes, K., and Vincent, K. (2006). Does tomorrow ever come? Disaster narrative and public perceptions of climate change. *Public Understanding of Science*, 15.

Lucas, G. (1977). Star Wars: Episode IV–A New Hope.

Lucian, o. S. (1902). *A true story*. A.H. Bullen.

Luokkala, B. (2014). *Exploring science through fiction*. Springer.

Léger, A., Rouan, D., Schneider, J., Barge, P., Fridlund, M., Samuel, B., Ollivier, M., Guenther, E., Deleuil, M., Deeg, H. J., Auvergne, M., Alonso, R., Aigrain, S., Alapini, A., Almenara, J. M., Baglin, A., Barbieri, M., Bruntt, H., Bordé, P., Bouchy, F., Cabrera, J., Catala, C., Carone, L., Carpano, S., Csizmadia, Sz., Dvorak, R., Erikson, A., Ferraz-Mello, S., Foing, B., Fressin, F., Gandolfi, D., Gillon, M., Gondoin, Ph., Grasset, O., Guillot, T., Hatzes,

A., Hébrard, G., Jorda, L., Lammer, H., Llebaria, A., Loeillet, B., Mayor, M., Mazeh, T., Moutou, C., Pätzold, M., Pont, F., Queloz, D., Rauer, H., Renner, S., Samadi, R., Shporer, A., Sotin, Ch., Tingley, B., Wuchterl, G., Adda, M., Agogu, P., Appourchaux, T., Ballans, H., Baron, P., Beaufort, T., Bellenger, R., Berlin, R., Bernardi, P., Blouin, D., Baudin, F., Bodin, P., Boisnard, L., Boit, L., Bonneau, F., Borzeix, S., Briet, R., Buey, J.-T., Butler, B., Cailleau, D., Cautain, R., Chabaud, P.-Y., Chaintreuil, S., Chiavassa, F., Costes, V., Cuna Parrho, V., De Oliveira Fialho, F., Decaudin, M., Defise, J.-M., Djalal, S., Epstein, G., Exil, G.-E., Fauré, C., Fenouillet, T., Gaboriaud, A., Gallic, A., Gamet, P., Gavalda, P., Grolleau, E., Gruneisen, R., Gueguen, L., Guis, V., Guivarc'h, V., Guterman, P., Hallouard, D., Hasiba, J., Heuripeau, F., Huntzinger, G., Hustaix, H., Imad, C., Imbert, C., Johlander, B., Jouret, M., Journoud, P., Karioty, F., Kerjean, L., Lafaille, V., Lafond, L., Lam-Trong, T., Landiech, P., Lapeyrere, V., Larqué, T., Laudet, P., Lautier, N., Lecann, H., Lefevre, L., Leruyet, B., Levacher, P., Magnan, A., Mazy, E., Mertens, F., Mesnager, J.-M., Meunier, J.-C., Michel, J.-P., Monjoin, W., Naudet, D., Nguyen-Kim, K., Orcesi, J.-L., Ottacher, H., Perez, R., Peter, G., Plasson, P., Plesseria, J.-Y., Pontet, B., Pradines, A., Quentin, C., Reynaud, J.-L., Rolland, G., Rollenhagen, F., Romagnan, R., Russ, N., Schmidt, R., Schwartz, N., Sebbag, I., Sedes, G., Smit, H., Steller, M. B., Sunter, W., Surace, C., Tello, M., Tiphène, D., Toulouse, P., Ulmer, B., Vandermarcq, O., Vergnault, E., Vuillemin, A., and Zanatta, P. (2009). Transiting exoplanets from the CoRoT space mission * - VIII. CoRoT-7b: the first super-Earth with measured radius. *Astronomy and Astrophysics*, 506(1):287–302.

Madhusudhan, N., Lee, K. K. M., and Mousis, O. (2012). A possible carbon-rich interior in Super-Earth 55 Cancri e. *The Astrophysical Journal Letters*, 759(2).

Mariscal, C. (2015). Universal biology: Assessing universality from a single example. In Dick, S., editor, *The Impact of Discovering Life Beyond Earth*, pages 113–126. Cambridge University Press.

Maunder, E. (1913). *Are the planets inhabited?* Harper and Brothers.

Mayor, M. and Queloz, D. (1995). A Jupiter-mass companion to a solar-type star. *Nature*, 378(2):355–359.

McGregor, G. (2020). Netflix enlisted 'Game of Thrones' team to adapt China's best-selling sci-fi novel. Then came controversy. *Fortune*.

McMahon, S. (2021). Astrobiology (overview). *Oxford Research Encyclopedias*.

Meadows, V., Reinhard, C., Arney, G., Parenteau, M., Schwieterman, E., Domagal-Goldman, S., Lincowski, A., Stepelfeldt, K., Rauer, H., DasSarma, S., Hegde, S., Narita, N., Deitrick, R., Lustig-Yaeger, J., Lyons, T., Siegler, N., and Grenfell, J. (2018). Exoplanet biosignatures: Understanding oxygen as a biosignature in the context of its environment. *Astrobiology*, 18(6).

Messeri, L. (2016). *Placing Outer Space: An Earthly Ethnography of Other Worlds*. Duke University Press.

Michel, J., Shen, Y., Aiden, A. P., Veres, A., Gray, M., Team, T. G. B., Pickett, J., Hoiberg, D., Clancy, D., Norvig, P., Orwant, J., Pinker, S., Nowak, M., and Aiden, E. (2011). Quantitative analysis of culture using millions of digitized books. *Science*, 331(6014):176–182.

Moretti, F. (2005). *Maps, graphs, trees*. Verso Books.

Morowitz, H. and Sagan, C. (1967). Life in the clouds of Venus? *Nature*, 215:1259–1260.

Mutonyi, H. (2015). Stories, proverbs, and anecdotes as scaffolds for learning science concepts. *Journal of Research in Science Teaching*, 53(6):943–971.

NASA Exoplanet Archive (2024). NASA exoplanet archive.

"NASA's Webb" (2022). NASA's Webb reveals steamy atmosphere of distant planet in detail. *NASA*.

Nash (2023). Box office history for science fiction.

National Academies Decadal Survey (2021). Pathways to discovery in astronomy and astrophysics for the 2020s. Technical report, National Academies of Sciences, Engineering, and Medicine.

National Research Council (2007). *The Limits of Organic Life in Planetary Systems*. The National Academies Press, Washington, DC.

Netburn, D. (2020). In the clouds of Venus, scientists may have found signs of extraterrestrial life. *The Los Angeles Times*.

Nováková, J., editor (2023). *Life beyond us*. Laksa Media and the European Astrobiology Institute.

Orthia, L. (2019). How does science fiction television shape fans' relationships to science? Results from a survey of 575 'Doctor Who' viewers. *JCOM*, 18(4).

Pearl, J. (1988). *Probabilistic reasoning in intelligent systems: networks of plausible inference*. Morgan Kaufmann.

Pichara, K. and Protopapas, P. (2013). Automatic classification of variable stars in catalogs with missing data. *The Astrophysical Journal*, 777(2).

PLATO Definition Study Report (2017). PLATO definition study report. Technical report, European Space Agency.

Propp, V. (1968). *Morphology of the fairy tale*. American Folklore Society and Indiana University.

Puranen, E. (2022). Through a mirror, earthly: Solaris, Gaia, and the search for other worlds. In E, G., editor, *Stanisław Lem and His Aliens: A Tribute and a Challenge*. Guardbridge Books.

Radić, V. (2021). The highest-selling games developed by BioWare ranked (and how much they sold). *Gamerant*.

Rauer, H., Catala, C., Aerts, C., Appourchaux, T., Benz, W., Brandeker, A., Christensen-Dalsgaard, J., Deleuil, M., Gizon, L., Goupil, M. J., Güdel, M., Janot-Pacheco, E., Mas-Hesse, M., Pagano, I., Piotto, G., Pollacco, D., Santos, Ċ., Smith, A., Suárez, J. C., Szabó, R., Udry, S., Adibekyan, V., Alibert, Y., Almenara, J. M., Amaro-Seoane, P., Eiff, M. A.-v., Asplund, M., Antonello, E., Barnes, S., Baudin, F., Belkacem, K., Bergemann, M., Bihain, G., Birch, A. C., Bonfils, X., Boisse, I., Bonomo, A. S., Borsa, F., Brandão, I. M., Brocato, E., Brun, S., Burleigh, M., Burston, R., Cabrera, J., Cassisi, S., Chaplin, W., Charpinet, S., Chiappini, C., Church, R. P., Csizmadia, S., Cunha, M., Damasso, M., Davies, M. B., Deeg, H. J., Díaz, R. F., Dreizler, S., Dreyer, C., Eggenberger, P., Ehrenreich, D., Eigmüller, P., Erikson, A., Farmer, R., Feltzing, S., de Oliveira Fialho, F., Figueira, P., Forveille, T., Fridlund, M., García, R. A., Giommi, P., Giuffrida, G., Godolt, M., Gomes da Silva, J., Granzer, T., Grenfell, J. L., Grotsch-Noels, A., Günther, E., Haswell, C. A., Hatzes, A. P., Hébrard, G., Hekker, S., Helled, R., Heng, K., Jenkins, J. M., Johansen, A., Khodachenko, M. L., Kislyakova, K. G., Kley, W., Kolb, U., Krivova, N., Kupka, F., Lammer, H., Lanza, A. F., Lebreton, Y., Magrin, D., Marcos-Arenal, P., Marrese, P. M., Marques, J. P., Martins, J., Mathis, S., Mathur, S., Messina, S., Miglio, A., Montalban, J., Montalto, M., Monteiro, M. J. P. F. G., Moradi, H., Moravveji, E., Mordasini, C., Morel, T., Mortier, A., Nascimbeni, V., Nelson, R. P., Nielsen, M. B., Noack, L., Norton, A. J., Ofir, A., Oshagh, M., Ouazzani, R. M., Pápics, P., Parro, V. C., Petit, P., Plez, B., Poretti, E., Quirrenbach, A., Ragazzoni, R., Raimondo, G., Rainer, M., Reese, D. R., Redmer, R., Reffert, S., Rojas-Ayala, B., Roxburgh, I. W., Salmon, S., Santerne, A., Schneider, J., Schou, J., Schuh, S., Schunker, H., Silva-Valio, A., Silvotti, R., Skillen, I., Snellen, I., Sohl, F., Sousa, S. G., Sozzetti, A., Stello, D., Strassmeier, K. G., Švanda, M., Szabó, G. M., Tkachenko, A., Valencia, D., Van Grootel, V., Vauclair, S. D., Ventura, P., Wagner, F. W., Walton, N. A., Weingrill, J., Werner, S. C., Wheatley, P. J., and Zwintz, K. (2014). The PLATO 2.0 mission. *Experimental Astronomy*, 38(1-2):249–330.

Reinsborough, M. (2017). Science fiction and science futures: considering the role of fictions in public engagement and science communication work. *JCOM*, 16(4).

Rimmer, P. B. (2023). *Origins of Life on Exoplanets*, chapter 16, pages 407–424. John Wiley & Sons, Ltd.

Robinson, K. (2015). *Aurora*. Orbit.

Sagan, C. (2010). Definitions of life. In Bedau, M. and Cleland, C., editors, *The Nature of Life: Classical and Contemporary Perspectives from Philosophy and Science*, page 303–306. Cambridge University Press.

Sagan, C. and Salpeter, E. (1976). Particles, environments, and possible ecologies in the Jovian atmosphere. *Astrophysical Journal Supplement Series*, 32:737–755.

Sagear, S. and Ballard, S. (2023). The orbital eccentricity distribution of planets orbiting M dwarfs. *Proceedings of the National Academy of Sciences*, 120(23):e2217398120.

Sanjin, S. (1908). *Japanese fairy tales*. A.L.Burt Company.

Saunders, D., Brake, M., Griffiths, M., and Thornton, R. (2004). Access, astronomy and science fiction: A case study in curriculum design. *Active Learning in Higher Education*, 5(1):27–42.

Schulze-Makuch, D. and Irwin, L. N. (2008). Ideas of exotic forms of life. In Schulze-Makuch, D. and Irwin, L. N., editors, *Life in the Universe: Expectations and Constraints*, pages 149–154. Springer Berlin Heidelberg.

Schwieterman, E., Kiang, N., Parenteau, M., Harman, C., DasSarma, S., Fisher, T., Arney, G., Hartnett, H., Reinhard, C., Olson, S., Meadows, V., Cockell, C., Walker, S., Grenfell, J., Hegde, S., Rugheimer, S., Hu, R., and Lyons, T. (2018). Exoplanet biosignatures: A review of remotely detectable signs of life. *Astrobiology*, 6:663–708.

"Science Consultant" (2019). Being a Hollywood science consultant isn't exactly glamorous. *Wired*.

Shapley, H. (1953). *Climatic change: evidence, causes, and effects*. Harvard University Press.

Shklovsky, V. (1925). *Theory of prose*. Dalkey Archive Press.

Slusser, G. (2003). The Ghost in the Canon: 'High Culture' Repressions of Science Fiction'. In Slusser, G. and Barricellie, J., editors, *Genre at the Crossroads: The Challenge of Fantasy*, pages 3–23. Xenox Books.

Slusser, G. and Chatelain, D. (2002). Conveying unknown worlds: patterns of communication in science fiction. *Science Fiction Studies*, 29(2):161–185.

Smith, V. A. (2010). Revealing Structure of Complex Biological Systems Using Bayesian Networks. In Estrada, E., editor, *Network Science*, pages 185–204. Springer.

Smithsonian Science (2011). From Star Wars to science fact: Tatooine-like planet discovered. *Smithsonian Insider*.

Snellen, I. A. G., Guzman-Ramirez, L., Hogerheijde, M. R., Hygate, A. P. S., and van der Tak, F. F. S. (2020). Re-analysis of the 267 GHz ALMA observations of Venus - no statistically significant detection of phosphine. *Astronomy & Astrophysics*, 644:L2.

Snow, C. (1959). *The Two Cultures*. Oxford University Press.

Stableford, B. (1999). *The Dictionary of Science Fiction Places*. The Wonderland Press.

Stableford, B. (2003). Science fiction before the genre. In James, E. and Mendlesohn, F., editors, *The Cambridge companion to science fiction*, pages 15–31. Cambridge University Press.

Stanway, E. (2022). Evidencing the interaction between science fiction enthusiasm and career aspirations in the UK astronomy community.

Stepney, S. (2015). The real science of science fiction. *The Guardian*.

Strickland, A. (2022). The weather on this exoplanet includes metal clouds and rain made of precious gems. *CNN*.

Suhner, L. (2017). The terminator. *Nature*, 542(512).

Suvin, D. (1979). *Metamorphoses of Science Fiction*. Yale University Press.

Szerszynski, B. (2019). Indigenous worlds and planetary futures. In Bold, R., editor, *Indigenous perceptions of the end of the world*, pages 203–210. Palgrave Macmillan.

Tasker, E. (2017). *The Planet Factory*. Bloomsbury Sigma.

Tattersdill, W. (2016). *Science, Fiction, and the Fin-de-Siècle Periodical Press*. Cambridge University Press.

Tinetti, G., Drossart, P., Eccleston, P., Hartogh, P., Heske, A., Leconte, J., Micela, G., Ollivier, M., Pilbratt, G., Puig, L., Turrini, D., Vandenbussche, B., Wolkenberg, P., Beaulieu, J.-P., Buchave, L., Ferus, M., Griffin, M., Guedel, M., Justtanont, K., and Zwart, F. (2018). A chemical survey of exoplanets with ARIEL. *Experimental Astronomy*, 46.

Tsiaras, A., Waldmann, I., Tinetti, G., Tennyson, J., and Yurchenko, S. (2019). Water vapour in the atmosphere of the habitable-zone eight-Earth-mass planet K2-18 b. *Nature Astronomy*, 3:1086–1091.

Turbet, M., Bolmont, E., Bourrier, V., Demory, B.-O., Leconte, J., Owen, J., and Wolf, E. (2020). A review of possible planetary atmospheres in the TRAPPIST-1 system. *Space Science Reviews*, 216.

UNSW Handbook (2024). Science and the Cinema.

UW Astrobiology (2024). Welcome to the University of Washington astrobiology program.

van de Kamp, P. (1969). Alternate dynamical analysis of Barnard's star. *Astronomical Journal*, 74:757–759.

Varela, F. J. (1997). Patterns of life: Intertwining identity and cognition. *Brain and Cognition*, 34(1):72–87.

Venables, M. (2013). Why Captain Kirk's call sparked a future tech revolution. *Forbes*.

Venturini, Julia, Guilera, Octavio M., Haldemann, Jonas, Ronco, María P., and Mordasini, Christoph (2020). The nature of the radius valley - hints from formation and evolution models. *Astronomy and Astrophysics*, 643.

Videla Rodriquez, E., Mitchell, J., and Smith, V. (2022). A Bayesian network structure learning approach to identify genes associated with stress in spleens of chickens. *Scientific Reports*.

Villanueva, G., Cordiner, M., Irwin, P., de Pater, I., Butler, B., Gurwell, M., Milam, S., Nixon, C., Luszcz-Cook, S., Wilson, C., Kofman, V., Liuzzi, G., Faggi, S., Fauchez, T., Lippi, M., Cosentino, R., Thelen, A., Moullet, A., Hartogh, P., Molter, E., Charnley, S., Arney, G., Mandell, A., Biver, N., Vandaele, A., de Kleer, K., and Kopparapu, R. (2021). No evidence

of phosphine in the atmosphere of Venus from independent analyses. *Nature Astronomy*, 5:631–635.

Vogogias, A., Kennedy, J., Archambault, D., Bach, B., Smith, V. A., and Currant, H. (2018). Bayespiles: Visualisation support for bayesian network structure learning. *ACM Transactions on Intelligent Systems Technology*, 10(1).

Voltaire (1752). *The works of Voltaire: Micromégas*. A. Fermin Didot.

von Trier, L. (2011). Melancholia.

Walker, G. (2012). The first high-precision radial velocity search for extra-solar planets. *New Astronomy Reviews*, 56(1):9–15.

Wolszczan, A. (1994). Confirmation of Earth-mass planets orbiting the millisecond pulsar PSR B1257+12. *Science*, 264(5158):538–542.

Wolszczan, A. and Frail, D. A. (1992). A planetary system around the millisecond pulsar PSR1257 + 12. *Nature*, 355:145–146.

Yu, J., Smith, V. A., Wang, P. P., Hartemink, A. J., and Jarvis, E. D. (2004). Advances to Bayesian network inference for generating causal networks from observational biological data. *Bioinformatics*, 20(18):3594–3603.