Cross-institutional MOOC Data Analysis and Visualisation: A call for collaboration

Manuel León-Urrutia  @mleonurr
Su White  @suuki
Lisa Harris @lisaharris

Adriana Wilde  @AdrianaGWilde
Graeme Earl  @GraemeEarl

FLAN – Glasgow, 19 April 2016
Outline

• Who are we?
• What do we do?
• What do we want?
• Will you join us?
Who are we?

Graeme Earl,
Lisa Harris (WSI)

Su White,
Adriana Wilde (ECS)

Manuel León-Urrutia
(ILIaD)

MOOC Observatory

http://blog.soton.ac.uk/mobs/
What do we do

• MOOC production
• MOOC data management
• MOOC data analysis
Our 12 published courses

... and more soon!
MOOC data management

An integrated toolset and data infrastructure: the MOOC Observatory and the UoS Web Observatory

MOOC Datasets in .csv

Converted to SQL

Converted to SQL

Stored & Hosted in UoS Web Observatory.

Queried & Analysed by MOBS
MOOC Data Analysis (some examples)

- Learner activity patterns
- Text Mining
- Social Network Analysis
- Real time visualisations
- ...

Learner Activity Patterns: Overall comments

(Number of comments in a MOOC against completion %)
Learner Activity Patterns: Comments per day

(Comments in a MOOC generated over time)
Learner Activity Patterns: Discussion generation analysis

(Number of comments that steps in a MOOC generate)
Learner Activity Patterns: Comparing course runs

Question answer patterns and clustering of failed responses
Evaluating impact of revised feedback content and mechanisms on results
Comparing assessment types and content e.g. peer review step text mining and analysis of stats

Comparing most commented steps on two runs:
Learner Activity Patterns: Comparing course runs

Two consecutive runs of the same MOOC exhibit similar behaviour (example: Sum of Steps Visited) when scaling over total steps.

Around week 7 there is an inflexion point.
Text mining: Portus MOOC comments

- Undertaking primary research about development and communication of archaeological knowledge (see next slide)

- Using concordance (AntConc), topic maps and other approaches to mine comments

- e.g. undertaking specific research such as examining the multisensory nature of creative writing on the course through co-occurrence of words (in this case “smell”)

<table>
<thead>
<tr>
<th>word</th>
<th>count</th>
</tr>
</thead>
<tbody>
<tr>
<td>fish</td>
<td>35</td>
</tr>
<tr>
<td>spices</td>
<td>29</td>
</tr>
<tr>
<td>sea</td>
<td>28</td>
</tr>
<tr>
<td>Sweat</td>
<td>18</td>
</tr>
<tr>
<td>Food</td>
<td>10</td>
</tr>
<tr>
<td>Cooking</td>
<td>8</td>
</tr>
<tr>
<td>animals</td>
<td>7</td>
</tr>
<tr>
<td>sewage</td>
<td>7</td>
</tr>
<tr>
<td>oil</td>
<td>5</td>
</tr>
<tr>
<td>fruit</td>
<td>4</td>
</tr>
<tr>
<td>smoke</td>
<td>4</td>
</tr>
<tr>
<td>wood</td>
<td>4</td>
</tr>
<tr>
<td>herbs</td>
<td>3</td>
</tr>
<tr>
<td>tar</td>
<td>3</td>
</tr>
<tr>
<td>awful</td>
<td>2</td>
</tr>
<tr>
<td>bread</td>
<td>2</td>
</tr>
<tr>
<td>damp</td>
<td>2</td>
</tr>
<tr>
<td>grass</td>
<td>2</td>
</tr>
<tr>
<td>baths</td>
<td>1</td>
</tr>
<tr>
<td>caulking</td>
<td>1</td>
</tr>
<tr>
<td>frankincense</td>
<td>1</td>
</tr>
<tr>
<td>garlic</td>
<td>1</td>
</tr>
<tr>
<td>grain</td>
<td>1</td>
</tr>
<tr>
<td>mud</td>
<td>1</td>
</tr>
<tr>
<td>olives</td>
<td>1</td>
</tr>
<tr>
<td>painting</td>
<td>1</td>
</tr>
<tr>
<td>Pitch</td>
<td>1</td>
</tr>
<tr>
<td>Plants</td>
<td>1</td>
</tr>
<tr>
<td>rubbish</td>
<td>1</td>
</tr>
<tr>
<td>sawdust</td>
<td>1</td>
</tr>
<tr>
<td>rubbish</td>
<td>1</td>
</tr>
<tr>
<td>seaweed</td>
<td>1</td>
</tr>
<tr>
<td>wine</td>
<td>1</td>
</tr>
</tbody>
</table>
Text mining: use of pronouns vs pedagogical value

Social Network Analysis: interactions between learners

SNA: networks of selected learners

Identification of highly connected individuals, divided into:

- those whose comments are replied to (receivers);
- And those who reply (givers)
Twitter MOOC Data Analysis

- Our MOOCs also generate data in other social media
- This example is from our Digital Marketing MOOC
- Learners were interacting with each other outside the Futurelearn Platform (Twitter in this case)
- We also have this dataset in our observatory
The MOOC Dashboard

Persuasive technologies

Fogg (2003) anticipated students in the future being persuaded to learn.

Vision: A student runs “StudyBuddy” on a hand-held device. Events:

1. she is congratulated on having met that day her daily study goal
2. she is presented suggestions on short, specific activities to engage with
3. she is presented a visualisation in which her peers who are also revising are represented in clusters as an encouragement
4. Her mentor can monitor her engagement and offers basic feedback.
Dashboards as persuasive technologies

The successful application of persuasive technologies in this context presuppose a very good understanding of the learners behaviour.

But this is challenging!

Data may be incomplete, inaccurate, technically difficult to collect and process in real time.
What do we want

• Share questions
• Test the dashboard
• Share data for cross-institutional analysis
• To (informally) know whether the network would want this
Now your turn!

• Please scan this QR code ...OR

• Go to http://respond.cc and enter 381871 as a session key