Crowdsourcing for Academic, Library and Museum Environments



This workshop, run as part of the Digital Humanities at Oxford Summer School 2015 (20 -24 July 2015), will enable participants to experience **crowdsourcing** in microcosm all the way from project conception to launch to data analysis. It will be a hands on and fast paced course, but there will be plenty of time to reflect on the process of setting up and sustaining a crowdsourcing project. Participants will come to the course prepared with a project idea and some sample data, e.g. 50 images of objects or books in their museum or library collection or an academic research dataset. It is absolutely essential that participants come prepared with their dataset and that this dataset is coherent. This means that the same string of questions can be applied to each image. For example, a dataset of medieval manuscripts might be processed in the following steps:

Are there any illuminations on this manuscript image? Yes/no;
If yes, draw a box around the illuminations;
A dropdown list titled ?What appears in the illumination?? Choices could include: animals, humans, plants, text, fantastical creatures, music etc.

On Monday participants will share and develop their ideas for their projects and hear from crowdsourcing experts, including ?Old Weather? Project Investigator **Philip Brohan**, of the Met Office. There will be the opportunity to ?storyboard? or ?wireframe? projects, so participants may want to print out a few images before arriving from their data, in order to draw on them and develop them.
By Tuesday they will have uploaded their data to **Zooniverse**'s new **Panoptes** DIY crowdsourcing site and launched a beta project that they will use for the rest of the week. Participants will have the opportunity to pitch their project to fellow course mates and try to generate interest in their project in order to gain experience in attracting a crowd and communicating the significance of their research or collection. By the end of the week the group will use data generated by their project (or back up sample data) in various data refinement and visualisation tools in order to learn the basics of how to manage and analyse their data.
This course will be of particular interest to **academics**, **librarians** and **museum professionals** who see the potential for crowdsourcing to expedite data extraction from non-machine readable collections. The Panoptes system will be particularly useful for metadata extraction projects and datasets that require a basic decision tree (yes/no answers and dropdown menus) but will not be able to support transcription at this time. Examples of the kinds of data extraction and workflows that will be supported include: www.penguinwatch.org and www.milkywayproject.org.
The workshop will be run by Dr **Victoria Van Hyning**, Digital Humanities Project Lead at **Zooniverse.org** (University of Oxford), and **Sarah de Haas**, a technical specialist from **Google** with a background in humanities who will help bridge the gap between humanities and technical skills.

The Digital Humanities at Oxford Summer School offers training to anyone with an interest in the **Digital Humanities**, including academics at all career stages, students, project managers, and people who work in **IT**, **libraries**, and **cultural heritage**. Delegates follow one of the 8 workshops throughout the week, supplementing their training with expert guest lectures. Bed and breakfast or

en-suite accommodation is available at St Anne's College, Oxford, on a first-come-first-served basis. Delegates can also join events each evening.

For further info visit http://dhoxss.humanities.ox.ac.uk/2015/crowdsourcing.html and http://dhoxss.humanities.ox.ac.uk/2015/ml/