

Engaging Crowds Advisory Board meeting minutes

Date of meeting: 17 February 2022

Attendees – Advisory Board

Stuart Dunn (SD)

Libby Elwood (LE)

Siobhan Leachman (SL)

Apologies – Advisory Board

Adam Corsini (AC)

Attendees – Engaging Crowds project

Samantha Blickhan (SB)

Elsbeth Haston (EH)

Sally King (SK)

Grant Miller (GM)

Bernard Ogden (BO)

Martin Salmon (MS)

Louise Seaward (LS)

Pip Willcox (PW)

Apologies – Engaging Crowds project

Chris Lintott, Zooniverse (CL)

Abbreviations:

RBGE Royal Botanic Garden Edinburgh

RMG Royal Museums Greenwich

TNA The National Archives

1.	Welcome, housekeeping and AOB suggestions Welcome from PW Apologies from CL and AC No suggestions for AOB
2.	Questions and clarifications from Report SL, LE and SD praised project team for their work

Team thanked SL for her report circulated in advance of meeting and AC for his comments sent by email
Project team will consider these comments for actioning in our next phase of work

SL noted an error in RMG's licensing conditions re: OGL and commercial reuse

MS thanked SL for alerting him

ACTION: MS to update licensing description on HMS NHS site (complete)

SL also noted licensing issues for RBGE

Reuse terms are in error or unclear on platforms like GBIF and Europeana
EH thanked SL – RBGE is aware of some of discrepancies and is working on getting them standardised

SL recognises that each institution is at different stages on their journey towards opening their collections

Board and project team discussed how licensing conditions can be restrictive at some museums and galleries – there can be tension between collections as revenue streams and making collections more accessible
SL comments and report can be fed into internal discussions

LE questioned whether it was possible in Zooniverse data to see if people used indexing tool purposefully or just went with first random page they found

GM explained that Zooniverse only records when someone submits a transcription

Tracking with Google Analytics not possible due to GDPR restrictions
So we can't follow exactly how people look at the records, as opposed to how they submit work on them

To study use of the indexing tool, we could download the data export, filter out sessions by user and then do a large-scale comparison of the first classifications submitted by everyone. This could give some indication of how volunteers are using the tool, as with these projects pages are otherwise served up sequentially. This is different from Zooniverse projects that do not use the indexing tool where subjects are served up at random.

BO noted that you could also look for evidence of people moving through the data non-linearly as they classify

PW noted that we will be asking for feedback on the indexing tool in our volunteer survey, which will be circulated soon

The 3 projects are slightly different

HMS NHS – selection options are limited to workflow and subject set
Scarlets and Blues and RBGE Herbarium – people can choose which page within a set they want to begin on and then work sequentially

3. **Discussion points**

Two discussion points were circulated to advisory board in advance

1. How appropriate is our plan for sharing data via the data sharing platform? Do you have any other considerations we should be thinking about?

Data will be shared via a page of project website
This will comprise shared datasets and links to images

SL suggested we give datasets a DOI and include licensing, citation information and a link to code we've developed

SB questioned whether we should make sure people also have access to the project interfaces – so people can look at these side by side with the data to understand the context in which the data was produced
SL suggested we archive each Zooniverse project – even GitHub may not be around in perpetuity

LE pointed out that educators who find the site might not be well-versed in academic terminology: we need to make sure descriptions are accessible to accommodate a spectrum of users

SD agreed that there is exciting potential to develop pedagogical exercises for students around these datasets

SL questioned whether there was a government data repository – there is one in New Zealand

There used to be an Arts and Humanities Data Service but this ended

There has been a recent initiative in Scotland: <https://research-scotland.ac.uk/>

AC said by email that he approved of the plans for the data sharing platform

He noted that if TNA's images move behind a paywall in the future, we should add an update to the webpage to explain the reasoning behind this

2. How can we make sure that the results of Engaging Crowds feed into the mission of Towards a National Collection (TaNC) to 'take the first steps towards creating a unified virtual 'national collection' by dissolving barriers between different collections – opening UK heritage to the world.?'

LE spoke about appreciating the concept of digital specimen – each document or object contains a huge network of data from different places that should be made available – context, literature, data, materiality etc. We need to embrace all angles of our records and enable people to make links for themselves as far as possible

EH noted that we are working towards the ambition step by step – it is very difficult and cannot be done quickly

SL said that a national collection must be able to be reused by the general public to do their own research

She encourages institutions to licence what will realistically generate income and open the rest of the collection as far as possible

If institutions relinquish some control, the public will step in and make new connections and discoveries

SD feels Engaging Crowds is challenging assumptions of what constitutes a barrier between collections

AHRC assumes that barriers between collections are institutional and technical

But this project is about breaking down barriers of user experience by exploring crowdsourcing across different collections

AC commented by email that it was important to promote and showcase this work as much as possible

	<p>This will help people across the sector realise how challenges can be surmounted</p> <p>This could be case studies but also events pitched on topics such as overcoming concerns about data sharing or transcription accuracy, and teaching people the first steps in creating a citizen research project</p>
4.	<p>Other responses to the Report: advice and suggestions</p> <p>None</p>
5.	<p>AOB</p> <p>PW thanked the project team for their ongoing work, and the advisory board for their enormous generosity in sharing their expertise and their time to guide the project's work</p> <p>Project team will consider advisory board's comments for actioning in our next phase of work</p>