

EnDOW: Enhancing access to 20th Century cultural heritage through Distributed Orphan Works clearance

Minutes of progress meeting 19th & 20th January 2017

WP0: Financial and Project Housekeeping

The deadline for the annual report was discussed. Maurizio noted that Bournemouth and Glasgow are currently reporting slightly over spend, but this was seen to be within tolerated amounts (under 10%).

Financial information from our colleagues at Bocconi University is required before 31st Jan. Marcella also stressed the need to send her updates for inclusion in the report: networking activity, dissemination, and related outputs. Simone reported her attendance at Forward event Brussels. Victoria noted her use of the EnDOW project during proceedings at WIPO in 2016.

Two academic papers are currently in progress. The 'Impossible quest' paper was reported by Simone to have been accepted, in progress toward a 2017 publication date. Lilla also has a paper submitted, no response yet from the journal.

Maarten raised the issue of the lack of budget for his work after June 2017. Approx. £15,000 will be required. Maurizio stated that the original aim was to re-allocate from different parts of budget. One solution is re-allocate part of IviR budget (subject to NVO approval).

Continuity planning: Simone announced her likely departure in August 2017-01-19 and Lucie will be moving to Canada effective July 2017-01-19. It was suggested that IviR will appoint a successor, or that duties may be split between current staff in IviR. Lucie will appoint a project controller and keep an eye on the budget from her new role. Lucie to advise on who will attend the progress meeting in Bournemouth in June on behalf of IviR

[KE Note: I have moved technical discussion into relevant work package sections below]

WP1: ASK Bocconi

Colleagues from ASK circulated a draft report on the national legal diligent search requirements.

Maurizio had a number of suggestions for the team about improvements to the draft paper. He wondered about the number of countries covered: the report should include 17 countries, but this number is inconsistent throughout the text. He suggested to make sure that summary reflects contents of the report. Also that definitions of accessibility should be added/made clear. Also we should consider the implication that the Directive is transposed

literally in most national legislations. There was a comment by Simone about the typology of sources and issues with how visual sources have been counted and presented.

Simone suggested using SPSS or another software tool for statistical analysis by coding variables, to enable other analyses and overall efficiency.

The team agreed timing of end of February for the next draft of the report incorporating changes.

Maurizio suggested that the WP1 and WP3 teams collaborate on two presentations for the June conference – one on legal and diligent search source data, and one on best practices by institutions (interview data).

Review of the Directive due in October 2016 from the Commission but they are not going to publish, other priorities. Suggestion that we should tell the commission we are analysing and we already know it's not working. When the report is finished, we should distribute and disseminate directly to selected policymakers. Contribute names on Basecamp if you think they are relevant.

Other milestone before July is the Best Practices report: for three countries, (or 3+1?) To show best practice in the sense of libraries and museums do in practice. Decided this will be done in conjunction with WP3 Glasgow interviews. Taking into consideration what the staff actually do. The groups discussed appropriate sampling strategy for this paper: How to find criteria for selecting institutions? Do we just look at ones we know are compliant, such as be searching on the EU database? Sectors and subject matter? Which sources are more relevant than others? How do you judge sources?

Fit was decided to interview 5 institutions from three countries: Italy, Netherlands and UK. Decided to focus on at least the 'big 3' national libraries and film archives in each country. First draft to be prepared by Bocconi/Glasgow by April, to have content to present at the conference on the 22nd June.

WP2: IViR Amsterdam

In the introduction to the meeting, Maurizio broadly raised the issue of improvements to the Diligentsearch.eu website, and in particular adding a 'resource' section or sub-page with more materials. One idea was to upload national information and flowcharts. There was debate about the purpose of this – e.g., If no one understands them, why put them online? There was also a suggestion to produce guidance for each country, such as booklets like the Out of Copyright calculators. It was decided that the present time is suitable to do error checking on the flowcharts, at the paper stage. The flowcharts are for practitioners, not lawyers. They're designed to be practical tools. It was decided that the flowcharts will go online with a short introduction to each one explaining the methodology. There was an additional question about credits: Do we have list of correspondents that contributed? Should they be credited somewhere on the site for the flowcharts?

We will proceed by:

- 1) Error-checking the Flowcharts (2 countries assigned to each investigator)
- 2) Convert to PDF
- 3) Maarten will tidy them up, add an introduction, add them to a new folder
- 4) Do 1-2 each to avoid going error-blind
- 5) License them so that people can adapt them

The deadline for the above is to be available by 22nd February for Maurizio to present at a meeting.

DOW platform

Maarten demonstrated the test version of the form builder and form on the EnDOW server. He stressed that the demo contained only bare bones tech features – no styling or designing at this stage. The code is available on Git Hub. Maarten estimated that the next version of the prototype would be ready from the developer in second week of February.

Needs to be restructured from current format to store the information that's entered. The output is exported as a PDF to be sent to the EUIPO.

Kris raised the question to the group about the lack of crowd-sourcing features in the current version of the proposed system. Specifically, the ability for the system to parse and distribute tasks out to various contributors. Members of the crowd need to be assigned work, and their completed work needs to be processed and collected. Do we have a plan for that separate system?

The response was that needs to be integration between the two, with the preferred development path to build it in little blocks so that lots of different institutions can use it. There is an API available to facilitate future additions/improvements of crowdsourcing.

Currently, each institution has to come up with a way of managing how the users access the form, access their collections and fill in the information. **Maarten raised the point that what we are offering currently is not crowdsourcing, but 'structured diligent search'.

Kris agrees but raises point that this has implications for work package 3 and evaluating DOW. Simulation – we will need to find people to pretend they are the crowd. They use a list of items and work their way through, we time them, observe them etc, then the results are sent back to the institution.

Maurizio suggested in practice what will happen is each Institution has their own community. They put a list of works online. They direct users to the web form.

There was a final discussion about translation of the web forms and system into different languages. The consensus was to create in English then translate elsewhere.

Another option was to use a plug-in for Google Translate. To facilitate translation, all text from the system needs to be in a spreadsheet, and translators complete each row, for easy switching between languages in the finished platform.

WP3: CREATE Glasgow

Kris presented on the methodology and progress of work package three. The presentation outlined the following aims:

Evaluating the eventual DOW tools

Informing prototype design of DOW through use-testing

Evaluate the platform's use and effectiveness across a variety of works:

Literary works, cinema and born-digital

Benchmark the benefits and costs for different types of works

The project milestones commitment is for two simulations – one pre DOW, one after DOW. We have completed simulation 1 on the Edwin Morgan collection. We will re-run the simulation using the DOW tool to see if it is more effective. Another way to evaluate is run through the system using institutions – observe what they do, how it works, how well, what they get back, how much they have to invest in it, participant observation

There was a discussion in general about the position of DOW among other technology proposals like blockchain. Can technology solve this problem? Given current initiatives at EU level – use of content ID, the creation of an OOC database? Opportunity to link the studies to these developments.***

There was speculation about where efficiency gains may arise with DOW. Kris is sceptical of gains from crowdsourcing (e.g. efficiencies of scale, since each worker must perform all of the same tasks, whether distributed or not). One potential savings is Time spent on writing the DS report – keeping a record of the search process – the tool does this for you, but the paper form doesn't, Kerry kept her records in a spreadsheet

Learning effect – do we expect archivists to get faster at entering information into the database? Is it better to have one person responsible for entering data vs a team, to take advantage of the learning effect?

Maurizio raised the question, might Collecting Societies interpret our calculations as a justification to charge excessive amounts for ECL licenses? (e.g. just below the amount we calculate it would cost to clear rights in orphan works for an institution).

We discussed the burdensome nature of the 'exhaustive list' of search databases in some legislations – this is major source of costs in our study. Germany have listed some sources in the implementation, but everything else is additional, but you are expected to search the sources in the legislation. Examples of institutions in Netherland that only search one database before making works available. Institutions that are worried and won't search until they are given explicit instructions.

Kris suggests that we should calculate the welfare effects of crowdsourcing: we have to assign a cost to the work the crowdsourcers are doing, but equally we have to consider search and bargaining costs of alternative schemes, like ECL. Simone added that we need to account for the motivation of crowd contributors – what do they get out of it? What are the social benefits of users doing it vs an archivist doing it internally – what do they learn? Do they feel like they're getting to contribute in a meaningful way? Victoria added the question about how much does taking risks cost institutions (not very much)?

It was decided to add unofficial ECL schemes to the interview script. Were you forced to take one? How much does it cost? What does it allow you to do? How many works are covered by it? Add question: Would they prefer an affordable ECL scheme? How much would you be willing to pay, - per work, per year etc. Add question – describe alternative legal solutions, including ECL, see which one they prefer.

Kris and Victoria proposed to incorporate these comments and circulate the script and design for final proofing by the team.

Closing Session and housekeeping

Arrangements for the June conference. It was decided that within the group we will be responsible for 3 presentations – One on general orphan works & endow plus platform demo; One on the countries study; One on the findings of the interviews. It was also suggested by Maurizio to include 2-3 case studies of orphan works being used – EM plus possible video games and something else. Mechanical Curator team could also present. Also suggested to ask Ben to speak about Spare Rib project. Perhaps the Belgian National Library, Lorna Hughes or Marissa from UCL or the Blockchain Australians if they can pay their fare. Consensus among the group that blockchain approach does not contribute meaningfully to orphan works problem.

Currently 2 U.S. speakers confirmed – Peter Jaszi and Meredith Jacobs. Pam Sam - if in Europe she might come and give a keynote. Or, the orphan works group at Berkeley might come over.

Title: ***New approaches to the Orphan Works Problem: Technology, Regulation and Practice***

Resources Page

Maurizio suggested content for the resources page: Create a separate short paper on the costs of diligent search from the EM article. Also create a separate short paper on what is diligent search based on Ronan's paper

Could the resource page be arranged by theme? What is a search? How much does it cost? Decision trees? Could we use a map? Also, should we present them separately as web pages rather than everything buried in PDFs that people won't read?

We discussed dissemination efforts beyond the website. Team members agreed to push the following in the next 5 months before June: Kluwer blog, CREATE blog, ASK blog, about the new report when it is published – IP Kat etc. Archive and Records Association, CILIP, Scottish Council on Archives. ASK to write the press release for the report when it is ready to go out.

Assignment of Tasks

Flowcharts have been assigned, they will be converted to PDF.
Agreed to complete by end of January 2017.

Annual report to go live by end of January once financial data is complete and everyone has updated their networking and dissemination activity – Marcella to nudge.

ASK report final draft to go live end of February – depending on response of final three outstanding questionnaires. Aura will follow up with the three countries who have not returned their questionnaires – the flowcharts can't be completed until this happens. Final for comment to be circulated, with feedback from Glasgow meeting included.

Victoria and Kris to share interview script with everyone for feedback, produce final version
Circulate first week of February once feedback is applied.

Next round of development on platform due to be completed by mid-February
PDFs to be made available in February in time for events – 10th February deadline

Information bubbles for the flowcharts – first drafts due end of March, final end of May
IViR to lead on this – they'll assign to a student to complete

Interviews UK to be completed by Victoria / Kris

Interviews Italy completed by Aura

Amsterdam contacts required – Victoria to do telephone interviews

Develop criteria then justify if we have to deviate – next week

Data collection should be completed by end of March, early April

Draft report on best practice due by end of May

Publication of June event programme end of February for promotion and invitations.