

Stories of Change Project
Principal Investigator: Joe Smith

AHRC Technical Plan

Section 1: Summary of Digital Outputs and Digital Technologies

The Stories digital strategy is to develop a repository, a repository of linked open data (RLOD), with a clean and inviting 'front end' web interface. In terms of functionality and graphic style this RLOD will draw upon examples such as www.bbc.co.uk (for clarity) and www.forensic-architecture.org (for the mix of research, outreach and impact functions). Examples of the front page, and one further sample page are presented as supplementary images with the Case for Support. We plan that over the course of the first two years:

- Content will flow from the three Stories and the synthesising Energetic work into the RLOD held on the OU's Knowledge Media Institute servers. Content will be uploaded by the project team and PDRAs
- All content (stories/interviews/historical data/literary quotes/pictures/film, in the form of text, audio, video and still images) is supplied by researchers (the team) with accompanying metadata (tags, keywords, paragraph of description, comment, references, links)
- The front of the RLOD website has a main 'hero item' of a slide show that scrolls through a selected body of our items of content (digital stories; interviews; pictures; quotes; artefact images; maps)
- All content of the RLOD will be publicly viewable from the moment it goes live (milestone: May 2013)
- A separate repository will allow anyone who registers to upload relevant content. This public repository will be reviewed monthly by the project manager, and any particularly relevant content will be transferred to the RLOD, with a 'flag' indicating that it originated as a publicly uploaded object.
- The website will automatically generate links to freely available relevant academic papers (via the existing online repository CORE tool) and learning materials (via OpenLearn), though these elements will be 'down the page' and supportive rather than prominent
- Engaging visualisations will present the content in compelling forms, both linear and networked
- Social media project activity, outputs and references (e.g. Twitter; FB; blogpost links) will appear instantly updated on a section of the page

At the beginning of year three the website landing page design will change. The RLOD will allow researchers to 'curate' a story through multiple content/data objects: viewers will be able to choose from a list of stories, and trace a path from one object to another, guided by explanatory text provided by the research team. As well as the curated story, alternative related items of content will be offered automatically as options by the RLOD. From this point onwards the home page will present three different ways of working with the material on the site:

1. Browsing the linked open data via visualisations or search terms / questions, pursuing self-generated routes through the data (by the research team or anyone).
2. Following one of e.g. half a dozen curated learning journeys through the data (primarily made for teachers and learners, and informal life-long learners; produced with intro level higher ed learners / higher school level in mind). These will be co-produced by the team as one of their outputs, in collaboration with project partners.
3. Following one of e.g. half a dozen more scholarly but approachable routes through the material. These would be short, concise and engaging digital versions of some of our full-length academic outputs.

The website will be maintained on the KMi servers for 3 years after the conclusion of the project, and all content will be archived offline by both KMi and the OU.

The design of our internet presence assumes that the route to the most effective and large scale public and policy uses of our repository of linked open data is to employ a social media editor part time across the life of the project whose prime responsibility will be to seed and grow diverse awareness and use of the resources. The editor will be proactive on social media platforms, responsive to shifting trends, and observant about the potential to place and promote content in relevant social media settings. We anticipate that they will work directly with project partners (including DECC and their commissioned communications agency Counter Culture). The social media editor will also support our face-to-face and mainstream media public engagement with carefully designed and targeted online events, collaborating particularly with PI J Smith and PDRA B Smith.

Section 2: Technical Methodology

2a: Standards and Formats

In total, we envisage that the Stories project will generate approximately 600 items of content (images, written documents, audio recordings, video clips, archive documents). This is based on an estimate that each of the three Stories and the 'Energetic' synthesising project will generate 150 items each. These content items will be uploaded to the repository by the project team.

However, the Stories repository will also hold material from previously conducted projects (such as 'What Happens When You Turn on a Light?') and generate new content through collaborations with project partners such as the Ashden awards. We will also invite other relevant projects funded by e.g. the AHRC and ESRC to input relevant content, We anticipate this will total a further 600 items. The repository will have the capacity to hold a very large number of items of content (60,000 items, or 100gb total data) to allow for such collaboration and expansion.

We envisage that each of the three Stories will produce approximately 15 hours of recorded interviews (audio). Interview recordings will be in mp3 format. These files are small so that downloading is easy; this will increase the accessibility of the open archive materials.

The Stories project will generate a relatively small number of video clips: approx 30 short clips (<10 mins), and 10 longer recordings (c. 30 mins) from, for example, workshops or seminars, and documentary or creative content produced by the team or partners. These will be in PAL format, 740x576 pixels at 25 fps, where possible, to ensure good quality.

The other content on the website will mostly take the form of documents and image files. Images will be at least 1280 pixel screen resolution, again to ensure quality.

2b: Hardware and Software

The website will use server space provided by KMi. The OU's web environment is load balanced and is hosting high traffic websites like OpenLearn and the OU's community website Platform. KMi has fully adequate hardware and software for this website, as well as the relevant expertise. The repository and associated website will be built with Drupal, as with other similar projects. The OU Drupal environment is currently hosting over 100 websites in multi-site installations. The OU's main IT unit is constantly monitoring the environment and is making sure that relevant security updates for hardware and software are implemented in a timely manner.

Interview material recorded by the research team will be edited using Audacity. This is a well-respected, free audio-editing software programme, easy to use and with a well-established online help community.

The PDRAs will need training in the use of Audacity audio editing software for lightly editing interviews for upload to the online repository. This training will be provided by The OU's LTS team, who have experience offering training on Audacity.

Following best practice for equal-access to online content, all of the audio and video content will be transcribed, and descriptions of visual content included.

2c: Data Acquisition, Processing, Analysis and Use

Design and construction of the repository will begin in January 2014, and a provisional webpage will be online by beginning March 2014. As design and construction proceeds, further functionality will be added to the repository and website. KMi will build the repository, in collaboration with a consultant responsible for graphic design and layout.

Content will be generated by the research team over the course of the project, starting with material from desk research from Jan 2014, and continuing with interview material and digital stories over the course of the project. Content will begin to flow into the online repository from March 2014, initially with help from KMi, and as functionality is added to the website, by PDRAs and the whole research team. Content will come from each of the three Stories, and from Energetic, the synthesising project. This content will include interview recordings and transcripts, 'digital stories', historical archive material, literary quotes and analysis, images and film clips, and other written documents resulting from desk-based research.

Details of all interviews, transcripts, audio and video recordings, and documentary archival work will be carefully recorded, and all content will be tagged with appropriate metadata when uploaded to the repository. This documentation will be maintained by the PDRAs, with backup copies held by KMi.

The repository will be open access from the start of the project. For the first two years of the project, the project team and project partners will work to populate the repository with content. A separate repository will also allow anyone who registers to upload items of content – these will be monitored and regularly reviewed by the project manager, and especially relevant and appropriate contributions will be transferred across to the Stories project repository, with a 'flag' to identify them as having originated as a public contribution.

A registration process (or similar process) will ensure that members of the public who wish to upload material are contactable by email to check rights or IP issues if necessary, and members will have to agree to a declaration that they hold the rights to any material that they upload to the repository. Advice on this process will be sought from the OU legal team to confirm that IP is being suitably protected.

Items of content in the repository will be linked through a combination of this metadata and semantic web technologies. All content will be searchable. The repository will allow the team to easily share all their research data with all other members of the team, allowing project team members to incorporate data from across the project into their research. Metadata and semantic web technologies will result in links made between disparate items of content, with innovative research implications.

Members of the research team will use the repository to curate 'paths' through the content, by linking a set of items together, and supplying supplementary explanatory text for each item. In year 3 of the project these paths will be shown on the website front page, and will be navigable by any visitor to the website. This work will draw on the tools developed for the DECIPHER project (concludes December 2013) designed by one of our digital team to support storytelling based on the collections of European museums.

Members of the public will also be able to curate their own paths through the material, though these will be identified as originating from public users of the website.

The research team will also upload to the repository, or link to, relevant academic outputs from the project as a whole.

All content on the repository will be regularly backed up by KMi.

Section 3: Technical Support and Relevant Experience

Advice has been sought primarily from Project Partner KMi, and in particular Zdenek Zdrahal. CI Zdrahal is bringing in his own experience of leading on the Bletchley Park Text digital humanities project and the EuroGene and CORE knowledge networking / semantic web projects. The Stories repository will be designed and built by KMi. Training in the use of the repository will be provided by KMi for the project team and project partners who will be using it.

Advice has also been sought from Bill Thompson (BBC), who will be a member of the Advisory Board. Thompson is leading the BBC's archive partnership work, and is a leading figure in web-broadcast-humanities interactions and a prominent technology journalist.

Internally advice has been sought from Prof Marie Gillespie (OU social sciences, PI on a major BBC World Service Trust project), and Andre Vock (Social Sciences Faculty Web Development).

Section 4: Preservation, Sustainability and Use

4a: Preserving Your Data

Advice regarding the preservation of data has been sought from the Social Science Faculty IT web office and KMi. The website will be maintained and monitored by KMi for 3 years after the end of the grant, as part of its ongoing maintenance program. This will include backing up the site, including all repository content.

Where appropriate consent has been received, interviews, images and other pieces of content will also be hosted on the Creative Climate website, for the lifetime of that project (2010-2020).

4b: Ensuring Continued Access and Use of Your Digital Outputs

The project's online repository is intended as a resource both for members of the general public with an interest in energy, low-carbon economies, climate change and related issues, as well as for those with more professional interests, for example teachers, students, members of the energy industries, local council officers and others. It will be metatagged appropriately so that it is easily found on common search engines such as Google. It will have a clean design for maximum impact and accessibility. It will be available directly for public access through its own website throughout the 3 years of the project.

After the end of the project, the repository of data will remain on KMi servers for 3 years (2017-2019), continuing to be accessible and usable by visitors to the site, though new uploads will no longer be possible. In the final 6 months of the project (i.e. from June 2016), KMi will advise the team on the viability of moving the repository to cloud computing (follow-on funding allowing), so that it can remain publicly accessible and usable beyond 2019.

Content hosted on the Creative Climate project pages of OpenLearn will remain searchable and publicly accessible and usable for the lifetime of that project (2010-2020).