A conference paper submission for OR2013 from the Digital Repository of Ireland

Title: Developing the Irish National Trusted Digital Repository for the Humanities and Social Sciences: an interdisciplinary approach.

Authors: Dr Natalie Harrower (RIA), Dr Sharon Webb (NUIM), Jimmy Tang (TCD), Damien Gallagher (NUIM), Eoin Kilfeather (DIT), Dr Seathrun Ó Tuairisg (NUIG), Dr Sandra Collins (RIA)

The Digital Repository of Ireland (DRI) is an interactive, trusted digital repository for social and cultural content held by Irish institutions. By providing a central internet access point and interactive multimedia tools, the DRI will facilitate engagement with contemporary and historical data, allowing the public, students, and scholars to research Ireland's cultural heritage and social life in ways never before possible. Since its initial conception, the DRI has taken a holistic approach to the practice of use and reuse, by targeting open source and community driven software, by carrying out extensive consultations across our interdisciplinary stakeholder community and by complementing the infrastructure and policy development with a comprehensive training programme to realise wide usage.

Launched in 2011, DRI is a four-year project funded by the Irish government's Programme for Research in Third-Level Institutions, based on a competitive bid. The DRI research consortium is comprised of six academic partners: Royal Irish Academy, National University of Ireland Maynooth, Trinity College Dublin, Dublin Institute of Technology, National University of Ireland Galway, and the National College of Art and Design. DRI is supported by and collaborates with a wide range of institutional stakeholders, such as the National Library of Ireland, the National Archives of Ireland, and the Irish national public service broadcaster, RTÉ.

DRI's position as a national trusted digital repository means that it must cater to and support a wide range of distinct and contrasting users, actors, and stakeholders. The requirements methodology is centered on qualitative interviews, dialogue and interaction between these various national partners, and has been formalised in comprehensive requirements statements. DRI's requirements elicitation phase extracted genuine cases of the use, reuse and repurposing of our partners' content as it exists presently (independent of DRI) but also

investigated how that same content may be linked together for richer use, reuse and repurposing given the capabilities of DRI's robust infrastructure. In 2012 DRI published a national report with the findings of its nationwide programme of stakeholder interviews to determine the digital preservation and access practices in cultural institutions, libraries, higher education institutions and funding agencies: http://goo.gl/O5CHf

The DRI is acting as a focal point for digital best practices by collaborating on the development of guidelines, and in 2013-2014, these guidelines will be consolidated and used to inform national policy on digital preservation and access. The DRI also promotes awareness of the benefits of preservation and open access to data, while respecting the importance of ownership, copyright, intellectual property, privacy and confidentiality. To this end, DRI has actively contributed to the National Steering Committee on Open Access Policy, and subsequently hosted the launch of the National Principles for Open Access Policy Statement by the Irish Minister for Research and Innovation in October 2012. (http://goo.gl/TLypK).

DRI supports the principle of "use, reuse and reproduce" by leveraging open source resources and community-driven software development. The adoption of open source software has had the positive effect of accelerating development and delivery of the DRI infrastructure. Informed by a broad study of available technologies, the project has chosen to make use of the Project Hydra infrastructure, based in turn on Fedora Commons and Apache SOLR. Further, an open source solution utilising commodity hardware will be selected to support activities in preservation, archiving, and dissemination of the project across multiple federated sites.

It became apparent through the requirements interview process that many Irish cultural institutions interested in collaborating with a national repository infrastructure had encountered barriers for digitising their artefacts or adapting existing digital projects into a long-term digital preservation cycle. As a result of this dialogue, DRI has elected to support a substantial variety of metadata and data standards; enabling DRI to use, reuse and adapt existing metadata, data, and digitisation protocols developed by these institutions and to provide the much-needed infrastructure for their long-term preservation. DRI is designed to support the use of Dublin Core, MODS, EAD, MARCXML and METS, all indexed so that they can be searched and faceted together through SOLR. This decision has been taken as a consequence of the processes of identifying the most commonly-used metadata formats in Ireland, researching the best practice metadata formats in digital repositories globally, and discovering the commonalities between the metadata fields, vocabularies and anthologies used in the selected metadata formats.

Continuing with its commitment to reuse, DRI is supporting and promoting the exploitation of its REST API to allow access to its data management and search systems. The provision of a well-documented API will allow educational, research and cultural institutes to provide new tools that would expand on DRI's core features. This principle encourages the rich use and reuse of content within the DRI repository in as open a manner as possible, while access control and authentication functionality will ensure any rights and confidentiality issues associated with the content are duly respected.

The core user interface work in DRI is building on the functional prototype using the Hydra-Head framework. A broad survey of digital asset management solutions from a user interface perspective has informed this aspect of the system requirements. A key requirement of the user interface is that it should provide a flexible suite of finding aids for cross-collection searches. Initially the tools will be implemented bilingually, in English and in Gaeilge (the Irish language), but building this framework allows for easy extension to other languages. The user interface design is intended to be responsive and provide access to a range of devices using HTML5 techniques. User involvement is an important part of the design process and design workshops in 2013 will include key stakeholders. The DRI user interfaces will include HTML5 multi-modal rendering technologies (e.g. animated media) streaming audio and video, (and where appropriate, 3D), and client side scripting (using the JQuery framework) for enhanced performance and user experience. These core interfaces will be supplemented with a digital curation tool, Storyscope, which has been developed by the sister project DECIPHER (http://decipher-research.eu/), funded by the European Union's 7th Framework Programme for Research and Technological Development. This flexible tool will utilise the DRI API to enable users to search the DRI collections, import digital objects and metadata, and organise these into meaningful narratives.

The DRI will be launched along with a series of five 'Demonstrator Projects' - digitally curated collections of unique content. Offering distinct depictions of various facets of Irish culture and society, these demonstrations will also serve several crucial functions during the developmental stage of the DRI infrastructure. Their successful development will ensure the reliability, scalability, and flexibility of the DRI to accommodate a broad range of resources, including video, audio, image and text. They will provide useful scenarios with which to test DRI's policies relating to data format and metadata standards, as well as how to manage multilingual content and how to deal with rights, ownership, ingest and access issues. The demonstrator projects will also be used to develop and pilot analytical and visualisation tools that will generate added value in analysing data. The projects will provide the DRI with an immediate critical mass of archival content that will be of interest to both researchers and the general public, and the demonstrators will also enhance relations with external stakeholders, who will both contribute to, and benefit from, many of these projects.

The DRI is also delivering a multi-faceted nationwide education and outreach program, aimed at bringing together members of the community, and enhancing digital skillsets for institutional staff, students, and stakeholders who produce, use or are responsible for digital content. The DRI will provide hands-on training for data producers and custodians preparing data for ingest, and is also delivering seminars and workshops on, for example, metadata and file format standards and approaches to linked data in an effort to promote creative and collaborative uses of repository materials. All the areas of DRI development – requirements gathering, repository architecture, added-value services, policy development and education – work in dialogue with one another, which ensures that we are building not only a multidisciplinary repository dedicated to use, reuse and repurposing, but also an interdisciplinary community ecosystem and nationwide awareness of digital preservation and access commensurate with these aims.

Contact details for follow-up;

Dr Sandra Collins

Director, Digital Repository of Ireland

Royal Irish Academy

19 Dawson Street

Dublin 2.

web: www.dri.ie

email: s.collins@ria.ie