

CURIOS: Connecting Community Heritage through Linked Data

Dr. David Beel
d.e.beel@abdn.ac.uk

Nophadol Jekjantuk
n.jekjantuk@abdn.ac.uk

Dr. Gemma Webster
gwebster@abdn.ac.uk

Stuart Taylor
staylor@abdn.ac.uk

Prof. Chris Mellish
c.mellish@abdn.ac.uk

Prof. Claire Wallace
c.wallace@abdn.ac.uk

dot.rural RCUK Digital Economy Research Hub
University of Aberdeen
AB24 3UE

ABSTRACT

The CURIOS project explores how digital archives for rural community heritage groups can be made more sustainable so that volunteer members can maintain a lasting digital presence. It is developing software tools to help remote rural communities to collaboratively maintain and present information about their cultural heritage. The objective is to investigate the use of semantic web/linked data technology to build a general, flexible and “future proof” software platform that could help such projects to develop digital archives and to be sustainable over time. As an interdisciplinary project we aim to synthesise a narrative that draws from both social science and computer science perspectives by critically reflecting upon the novel approach taken and the ongoing results that are being produced.

Categories and Subject Descriptors

E.2 [Data]: Data Storage Representations – *Linked representations*, J.4 [Computer Applications]: Social and Behavioral Sciences – *Sociology*, K.4.0 [Computers and Society]: General.

General Terms

Human Factors

Keyword

Cultural Heritage, Digital Archives, Community heritage, Open linked data.

1. MOVING TOWARDS COMMUNITY DIGITAL HERITAGE

Rural areas are characterised by a strong identity of people with place. These identities draw on a repertoire of cultural norms, knowledge, histories, customs and practices which, taken together, construct unique place identities. This cultural distinctiveness is dynamic given traditional cultural practices are reproduced and others introduced as cultural systems evolve and adapt. Forms of cultural expression, such as story-telling, music

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

Conference '10, Month 1–2, 2010, City, State, Country.
Copyright 2010 ACM 1-58113-000-0/00/0010 ...\$15.00.

and song, poetry and literature, dance and drama together with material objects, artefacts, sites and cultural spaces, are resources for interacting with the past and for experiencing the present. In the collection and transmission of these collections there has been a growing sense that the traditional methods for doing this are failing, Nora [1]. In order to address this problem, digital solutions have been sought but this has been a problematic process due to a number of variances. These include the constant changing of file types, software and codes of best practice, as well problems to do with cost and the sheer amounts of ‘analogue’ data to convert. Leading the way in this process have been national institutions but with the production of such local cultural repertoires, which as Flynn [2] suggests ‘*are the grassroots activities*’ where ‘*control and ownership of the project is essential*’ there has been a failure to consider the needs of community heritage groups in these processes. As such groups do not want to be subsumed into national archives, which they do not control, is not sensitive to their needs and is juxtaposed ideologically to the production of their own ‘place history’. Following Creswell’s [3] claim that such archives represent ‘spaces of marginalized memory’ CURIOS is therefore seeking a solution using open linked data in which a system can be developed that is attuned to the specificity of a local heritage but can also take advantage of already collected materials from elsewhere.

2. CASE STUDY – HEBRIDEAN CONNECTIONS

In the past 40 years around 22 ‘Comainn Eachdraidh’¹ (CE), have been established in the Outer Hebrides². CE are community run groups that began in the 1970’s with a very specific political and cultural purpose – to preserve the culture, history and language of the primarily Gaelic regions of Scotland. Such community heritage practices have been described as a ‘messy’ endeavour with a wide variety of different formal and informal practices [5]. The archives embrace different registers of social memory from tangible to intangible heritage, which have been collected and ordered in a variety of different ways. Different CE groups collect and order their archives in a variety of different ways: from the highly ‘professional’ to the more bespoke and sporadic. As the CE groups are voluntary community archives, they are rooted in local

¹ Comainn Eachdraidh is a Gaelic phrase meaning ‘Historical Society’.

historical values, hence there is often little consistency between groups regarding cataloguing, archiving and content management.

Hebridean Connections (HC), which is a community managed, online historical resource was formed due to the driving force of a single member of a CE who saw the benefit of digitising and connecting the different historical catalogues [5]. The idea was proposed to the different CE and four groups were actively involved in a Heritage Lottery Fund (HLF) bid that funded the creation of the HC website³. The project website was launched in 2006, holding some 100,000 records relating to the genealogy, history, archaeology, and cultural traditions of the Outer Hebrides. Currently, the system allows users to search using keywords, selecting relevant images, or with a map-based interface. Additionally, the website encourages contributions from its users and, therefore, has the potential to foster reciprocal knowledge exchange across geographical boundaries.

2.1 Sustainability

HC is one example of a community-built digital cultural heritage repository where their long-term future is unclear. Many issues with the current system have arisen since the initial grant, particularly surrounding funding and scalability. There is a real practical question about how a project of this kind can be maintained over time with the resources available to a small dispersed community, especially as the initial system was developed by a private development company, using proprietary software. As the project developed, this situation raised the problem that any changes to the system required more financial investment in the software. For the small community heritage groups involved, this was not feasible, especially as the CE became aware of what was possible through digitisation and wanted to expand. The process of digitisation has created two primary issues for HC:

- How to expand the project remit without additional funding for developers?
- Scalability issues, how can more CE collections be integrated in a closed system?

2.2 An Archive for the Future?

Motivated by the limitations of the current HC system, the CURIOS project's aim is to produce a sustainable system that allows a community of users to manage a digital archive of cultural heritage data, or 'cultural repository', releasing them from any specific proprietary software platform. To achieve this goal, CURIOS has made use of existing open source content management system (CMS) software and Semantic Web standards.

The emergence of the Semantic Web [6] has led to several standard formats for representing and interchanging data [7, 8]. In 2009, Tim Berners-Lee proposed linked data as a set of best practices using standard web technologies to publish and link data on the web [9]. This has led to a vast increase in the number of linked data datasets becoming available on the web [10]. By making use of linked data, cultural repositories would have the

potential for reuse and integration with further related data sources.

In recent years content management systems have gained popularity on the web by allowing users to build and publish web pages without requiring in-depth knowledge of the underlying web technologies. The CURIOS project has extended the web CMS approach to allow users to manage repositories of linked data. This Linked Data CMS approach makes use of existing CMS software to retain the usability and scalability of existing tools that are familiar to users, whilst allowing the users to exploit the benefits of linked data.

The Linked Data CMS approach has been implemented as a module for the popular open source Drupal CMS⁴. Building the next generation of Hebridean Connections on open source software and web standards has distinct advantages for future development and use of the system. The Drupal-based system can be maintained by its community of users and can be extended additional functionality developed by the Drupal open source community, e.g., to support blogging or e-commerce features. This community led maintenance allows for further future development of the cultural repositories as the archives develop.

2.3 On-going Challenges

There are a number of on-going challenges in developing a sustainable user driven cultural repository. This is the focus of research questions for CURIOS currently. Firstly, as new cultural heritage vocabularies and datasets emerge, how can users be encouraged to integrate the Hebridean Connections cultural repository with these datasets and vocabularies? Secondly, the Drupal-based system is currently tailored to the cultural heritage domain. How should the approach be generalised so that it can be released to the open source community and applied to a wider variety of domains on the web? Finally, to investigate how semantic web reasoning technologies can be applied in a scalable way to automatically extract the implicit information from these cultural repositories?

3. CONCLUSION

Open linked data can help make local cultural repositories sustainable and collective. Linked data allows for collaboration, mutual authoring, distributed responsibilities through community projects and the utilisation of other community or national resources [4]. The CURIOS project is enabling local cultural heritage repositories to become a meaningful identity resource for an international community, who previously had no access to them. By falling outside of national institutional frameworks, local people are the 'gatekeepers' of their own heritage and are selecting what to commemorate based on their own customs of remembering. This kind of digital archive can have, therefore, potentially significant social impacts which need to be better understood. The vision of Hebridean Connections is to expand the collections to incorporate those held by other Comann Eachdraidh. Additionally, by making use of linked data, there is now the possibility to integrate further sources of data into HC from other historical societies or even national organisations.

² The Outer Hebrides is a group of islands off the West coast of mainland Scotland.

³ The Hebridean Connections website is hosted at <http://www.hebrideanconnections.com/>.

⁴ Drupal is a popular open source web content management system: <http://drupal.org/>.

4. ACKNOWLEDGMENTS

We would like to thank Hebridean Connections and the Comainn Eachdraidh for their ongoing commitment to this research. This work is supported by the Rural Digital Economy Research Hub (EPSRC EP/G066051/1).

5. REFERENCES

- [1] Nora, P. 1996. *Realms of memory: rethinking the French past. Volume 1: Conflicts and Divisions*. Columbia: University Press.
- [2] Flinn, A. 2007. Community Histories, Community Archives: Some Opportunities and Challenges 1 in *Journal of the Society of Archivists*. Volume 28, Issue 2.
- [3] Creswell, T. 2012 Value, gleaning and the archive at Maxwell Street, Chicago. *Transactions of the Institute of British Geographers*. Vol. 37 (1),1-13.
- [4] Mellish, C., Wallace, C., Tait, E., Hunter, C., & MacLeod, M. 2011. Can Digital Technologies increase Engagement with Community History? *Digital Engagement 2011*. <http://de2011.computing.dundee.ac.uk/wp-content/uploads/2011/10/Can-Digital-Technologies-increase-Engagement-with-Community-History.pdf>
- [5] Wallace, C., Tait, E., MacLeod, M., Mellish, C., & Hunter, C. 2011. Supporting Digital Humanities Creating Sustainable Digital Community Heritage Resources Using Linked Data. In *Supporting Digital Humanities: Answering theunaskable Conference*. 17–18.
- [6] Spector, A. Z. 1989. Achieving application requirements. In *Distributed Systems*, S. Mullender, Ed. ACM Press Frontier Series. ACM, New York, NY, 19-33. DOI= <http://doi.acm.org/10.1145/90417.90738>.
- [7] Berners-Lee, T., Hendler, J., and Lassila. O. 2001.The Semantic Web. *Scientific American*, 284(5), 34–43
- [8] Hitzler, P., Krötzsch, M., Parsia, B., Patel-Schneider, P.F., and Rudolph, S. 2009. (eds.) *OWL 2 Web Ontology Language: Primer*. W3C Recommendation, <http://www.w3.org/TR/owl2-primer/>.
- [9] Manola, F. and Miller. E. 2004. *RDF Primer*, W3C Recommendation, <http://www.w3.org/TR/rdf-primer/>.
- [10] Berners-Lee,T. *Linked-data design issues*. W3C design issue document, June 2009. <http://www.w3.org/DesignIssue/LinkedData.html>.
- [11] Bizer, C., Heath, T., and Berners-Lee. T. 2009. Linked data - the story so far. *International Journal of Semantic Web Information Systems*. 5(3), 1–22