

Infrastructures for Digital Research: New opportunities and challenges

*Lorna Hughes**

Introduction: understanding the use of digital content

The past twenty years has seen a very visible expansion of the digitization of Europe's Cultural Heritage. However, to put the scale of this investment in perspective, the ENUMERATE¹ project has produced data about the volume of digital cultural heritage available across Europe. The results of their analysis published in 2014² shows that over 10% of the collections of European museums, archives and libraries has been digitized, over 300 million items. At the present scale of progress, it will take over 30 years to digitize the rest: a task that will be more complex taking into consideration that a large amount of material remaining to be digitized is 20th and 21st century material, either moving image, audio, or born-digital. A growing number of organisations have developed digital strategies to address this: 36% of the institutions surveyed by ENNUMERATE in 2014 have a written digitisation strategy.

Of the content that has been digitized, though, only 34% is available online, and of that only 3% is available for open re-use, though Creative Commons and unrestricted licensing. However, more than half of the organisations surveyed do not implement measures to quantify how their digital content is used. This lack of analysis of use isn't just an issue in cultural heritage digitization: in 2008, the UK's Arts and Humanities Research Council carried out an analysis of sustainability

* Humanities Advanced Technology and Information Institute University of Glasgow.
Email. Lorna.Hughes@glasgow.ac.uk

¹ <http://www.enumerate.eu/>

² ENUMERATE 2014, 6.

planning for projects funded through its Resource Enhancement Scheme, 173 projects, an investment of approximately £40 million¹. Of these projects, only just over half were collecting usage statistics². A lack of enquiry into the use and impact of digital content is not a new issue: indeed, there has been a disproportionate investment in creation, management and curation of digital resources versus *use* of digital content for scholarship (in the UK from 1999-2009, the AHRC invested approximately £1.5 million into research that addressed the *use* of digital content, tools, and methods for research, but during this period they invested significantly more in digital content creation and curation). The lack of investigation in this area raises some serious questions about the value of the significant international investment in digital cultural heritage.

What are we doing with all this digital stuff?

It is perplexing that more institutions don't try to develop a better understanding of what their users do with digital content, because methodologies exist for gathering and analyzing this information. Using analytic tools including Google Analytics³ as a baseline, it is possible to augment numeric data with other qualitative and quantitative measures for assessing users, including structured surveys, interviews, and content analysis tools (looking at citations of digital content, and its embedding in blogs and other resources). These approaches have been refined in a tried and tested approach developed by the Oxford Internet Institute (OII), called Tools for the Impact of Digitised Scholarly Resources (TIDSR)⁴. This has been used successfully to measure the impact of a range of resources in the arts and humanities⁵. Similarly, methods developed by the CIBER initiative⁶, originally based at University College London, enable an understanding of the 'digital footprint' of users of heritage content.⁷

¹ Robey 2008.

² Robey 2011, 153

³ <https://www.google.co.uk/analytics/>

⁴ <http://microsites.oii.ox.ac.uk/tidsr/>

⁵ For an overview of the use of this method on a collection of Northern Irish Parliamentary Papers, see Hughes et al. 2015.

⁶ <http://ciber-research.eu/>

⁷ <https://www.ucl.ac.uk/news/news-articles/1001/10010802>, and see also Nicholas 2014.

These approaches are an interesting way to build a narrative around data: using statistics about use to build a picture of engagement with digital cultural heritage. Overall a striking finding is the change in information seeking behaviours: users will multitask, and do many things lightly rather than one thing deeply: they have become what David Nicholas of CIBER referred to as ‘web foxes’⁸ as people bounce around the digital domain using fast and abbreviated searches. Visitors to sites will frequently only access one page, then move on. This is even more noticeable when users access heritage data using mobile phones devices on the go: searching is shorter, and less engaged.

In order to investigate these questions in more detail, from 2011-14 I undertook some analysis of the use of the collections of the National Library of Wales, specifically *Welsh Journals Online* (<http://welshjournals.llgc.org.uk/>) and *Rhyfel Byd 1914-1918 a'r profiad Cymreig / Welsh experience of the First World War 1914-1918* (www.cymru1914.org)⁹. This analysis showed several interesting things. Google analytic data for both sites showed that users mostly come from search engines, not the library website or interface, which is interesting in terms of the investment in presenting content. Similarly, many users were referred through family history websites, or media stories about the resource, rather than academic or subject-specific resources that linked to it: the highest number of referrals to *Rhyfel Byd 1914-1918 a'r profiad Cymreig / Welsh experience of the First World War 1914-1918* came from a story on the BBC Wales News Website about the launch of the resource¹⁰. The most popular search results are location based: people are looking for historic information about places they know. Interviews with selected groups of users of both resources expanded on these findings, establishing that the search box enabled most searches: that the ‘googlization’ of information seeking had not affected the quality of information they were able to find. In fact, the only limitation that users reported was that there was not enough digital content in each resource. This is a consistent finding when surveying users of digital cultural heritage: as fast as digital content is launched, the demand increases for more

⁸ Nichols 2015, 23.

⁹ Welsh Journals Online: <http://welshjournals.llgc.org.uk/>. The study into the use of Welsh Journals Online was written about in Hughes 2012; the research into the use of [Cymru1914.org](http://www.cymru1914.org) was presented at the Sheffield Digital Humanities Conference, 2014: <http://www.hrionline.ac.uk/dhc/paper/19>.

¹⁰ <http://www.bbc.co.uk/news/uk-wales-25126781>.

material: 'we have too many digital collections' said no survey respondent, ever. The benefits of unlimited access to information are so great that they compensate for problems that arise from irrelevant data, poor metadata, or digital dead ends.

"We are all digital humanists now"

Humanities research also exemplifies this enthusiastic embrace of the digital: In many respects, we are all digital humanists now: we all use digital data for research by default (at the very least, electronic catalogues that point to analogue resources). Scholars rely on a steady consumption of digital source materials for scholarship and pedagogy, mostly created through large-scale digitisation initiatives in universities, libraries, museums and archives and as well as by commercial entities¹¹. And humanists do not just use data: many scholars now create and manage these resources; we communicate via blogs and twitter, and use digital dissemination methods for sharing and challenging research results. For many, this represents a sea change in research practice within the recognizable lifetime of a career.

But digital humanities is more than just the use of digital content for searching and browsing. If I was to attempt a definition of '*digital humanities*' as an emerging field of research, I would suggest that it is combination of using digital content; with digital methods for the analysis and interpretation of this content; tools for specific scholarly tasks; and communicating research to the widest possible audience using traditional and non-traditional publishing methods. At its most effective, digitally enabled research in the humanities can facilitate and enhance existing research, making research processes easier via the use of computational tools and methods. More interestingly, it can enable research that would be impossible to undertake without digital resources and methods, allowing the formulation of new research questions that are driven by insights only achievable through the use of new tools and methods.

At the core of this sort of approach is collaboration: with researchers from scientific and other humanities disciplines, computational and technical fields, as well as cultural heritage organizations.

¹¹ Ell and Hughes 2013, 24.

Another important collaboration is with the audiences for our work, as user-led design, and participation via mechanisms like crowdsourcing, informs the development of many initiatives. It also relies on underlying technical infrastructures.

I like this characterization of digital humanities as about *content, tools and methods*, because it creates a working environment with raw materials, tools for working with the raw material, and expertise in digital methods. This may seem a reduction of humanities practice, one that confirms Gary Hall's view of the digital humanities as a use of tools and technologies that does not rely on a theoretical framework¹²: the sort of model that one can see as the basis for a DH research infrastructure, putting the maker's perspective at the centre. However, another school of thought in the digital humanities, exemplified by Alan Liu, sees the emphasis on building and making as a space for exploring and critique of culture¹³. I would argue that the maker and theorist perspectives are not incompatible: in fact, it is through developing and building digital projects in the humanities *through practice* that we can conduct cultural and critical analysis, by questioning many of the assumptions on which digital resources are built and communicated, and developing a better framework for understanding the ways that working with digital content, tools and methods is transforming our consumption and production of knowledge. By subsequently analyzing use of the digital collections and content that have been built, we can also develop a better understanding of their role in the humanities research lifecycle, and start to address questions about how digital content in particular is not just helping us to do research more effectively, but a disruptive, transformative intervention in the research lifecycle.

To address some of these questions, from 2011-15 the European Science Foundation funded the Research Network in Digital Methods in the Arts and Humanities (NeDiMAH.eu), to look at the practice of digital humanities across Europe, and to understand what researchers need in order to do digitally enabled research in the future. From a detailed programme of methodologically focused activities, several conclusions can be drawn about what can

¹² Hall 2012.

¹³ Liu 2012.

make digital resources and collections more valuable for scholarship¹⁴. These conclusions sit alongside the user analysis described above to suggest three core observations about what the research community needs from digital content, and the infrastructures that deliver this content: access to, and ‘knowability’ of digital content; opening up content for use and re-use; and creating better environments for drawing together multimedia content from a variety of sources for analysis and publication. These are described in more detail in the sections that follow.

Access to, and ‘knowability’ of digital content

Much of what is classified as ‘digital humanities’ is consumption: scholarly use of digital resources and born-digital material, working in what Roy Rosenzweig called a culture of abundance.¹⁵

While this presents new possibilities it also raises significant challenges, and there is a need for a deeper understanding of the digital resources scholars rely on. As Tim Hitchcock has described in his article *‘confronting the digital: or how academic history writing lost the plot’*¹⁶, researchers often work with data that is what we would charitably call ‘limited’ in its potential for re-use, thanks to problems with OCR, markup, and description and provenance information. Similarly, limitations in metadata often mean that while keyword searching gives the satisfying ‘quick hit’ of a result, the user frequently misses important contextual information that enables more mindful engagement with sources, especially archival content: the scale of the original sources, their condition and their environment are often hidden when using digital surrogates that often appear as disembodied objects, separated from their context. Ryan Cordell, who has theorized a “network-author” function in antebellum newspapers from research based on digital newspaper archives, has said that “most digital archives hide more

¹⁴ Full reports of all NeDiMAH activities are available at the Network’s website, nedimah.eu. These activities were structured around Working Groups, each addressing a specific area of Digital Humanities methods: Space and Time; Information Visualization; Linked Data and Ontological Methods; Developing Digital Data; Large-Scale Text Collections; and Scholarly Digital Editions.

¹⁵ Rosenzweig 2010, 22.

¹⁶ Hitchcock 2013.

than they reveal, as keyword searches require prior knowledge of the texts to be discovered and can lead to evidentiary excess.”¹⁷

Another issue is the selection of content for digitization. Due to limitations of funding, availability of analogue sources, and copyright and licensing issues, many digital archives are not ‘complete’. The National Library of Wales received funding from the Welsh Government and European Regional Development Fund (ERDF) to digitize 1 million pages of Welsh Newspapers Online (<http://newspapers.library.wales/>) from its holdings. These date from 1800-1919, in Welsh and English. As a selection of the newspaper coverage of the time, this is an important resource for research and teaching: however, it is by no means a full representation of all newspapers of the period.

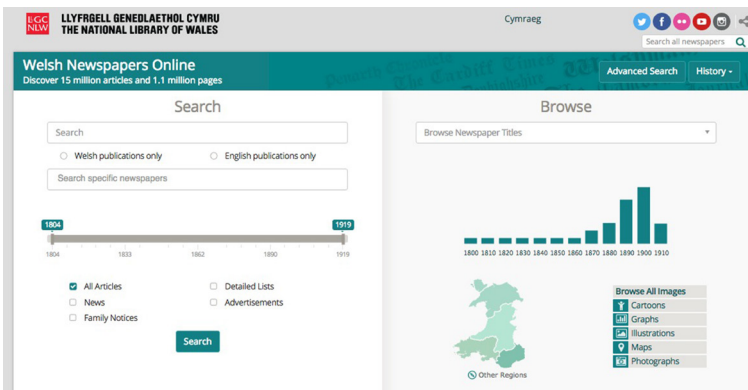


Fig. 1. Home page: Welsh Newspapers Online.

Figure 1 shows a chart on the home page of the resource showing the materials selected for digitization: many are missing from the 1909-19 period as a smaller, separate funding source was used for this. From the earlier period, many of the newspaper runs held by the library are incomplete. However, for the user, it’s too easy to assume that the resource is a complete representation of *all* Welsh Newspapers, ever: the chart only shows the proportion of the newspapers in the resource, not as a proportion of those published. As we know, users (especially students) frequently turn to digital content by

¹⁷ Cordell 2015.

default when looking for historical sources, and as a free resource, this newspaper archive is very likely to be used, leading to the likelihood of research being conducted from an incomplete dataset.

Opening up content for use and re-use

A second requirement for researchers is the need for better, and more open, solutions for analysis and linking of digital content, which is all too frequently locked in digital silos, unable to be easily re-used and linked with other data from other providers and collections. Evaluation of users of digital content, and assessment of a range of European Digital Humanities projects, shows that scholars frequently have very simple questions or ideas that they want to test with data at the desktop, and they do not want the technology to be a barrier. When data is locked in 'silos', it can be difficult for the end user to integrate the tools they need analysis and linking with content held remotely. There needs to be greater disassociation of text and data from platform and delivery mechanisms, liberating digital resources for purposes unanticipated by the creators of digital resources¹⁸.

For example, users often want to be able to use simple, pattern-finding tools, like nGram, with a range of data sets to quickly test hypotheses – the example in Figure 2 shows a visualization, using nGram, of the term 'Belgian Refugees' in newspapers in Welsh Newspapers Online from 1914-1919. The graph shows a spike of heightened interest in the almost fashionable cause of Belgian refugees in late 1914, which tailed off as the war continued. A similar pattern can be seen in newspapers from around the British Isles¹⁹. For questions like this, requiring simple pattern finding, this sort of tool can be a compelling and timesaving aid for researchers. However, as the tool is not part of the resource, using it with this collection required liaison with the National Library of Wales developer team to extract and work with the core data in this way.

One of the huge advantages of digital access is making the invisible visible in archives. We lose this when we can't easily re-use and re-purpose this content.

¹⁸ Robinson 2014, 251.

¹⁹ Hughes 2016.

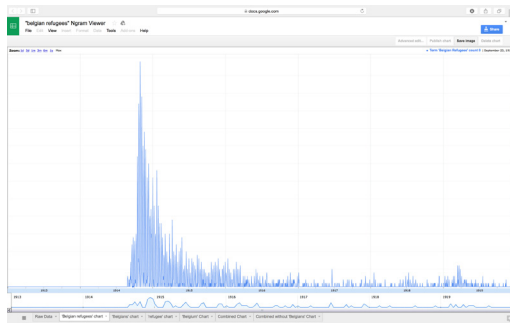


Fig. 2. nGram of the term 'Belgian refugees' in Welsh Newspapers Online.

Creating better environments for drawing together multimedia content from a variety of sources for analysis and publication

The attraction of working in a digital environment is the ability to integrate sources in a variety of formats, and to use them for new and unforeseen purposes. An example of working with hybrid, multimedia content can be seen in an AHRC funded project, *The Snows of Yesteryear: Narrating Extreme Weather*.²⁰ The project was a creative collaboration to uncover archives in the collections of the National Library of Wales that contain information about the impact of extreme weather in Wales during the period from about 1700 (the pre-weather instrument period) to the 1960s. This included manuscripts (literary works, diaries, letters); printed materials, including Welsh ballads; and other material, including art works and manuscripts. These were digitized and made accessible to a team of climate scientists from the ACRE Project at the UK Met Office who used the data to fill gaps in the global picture of extreme weather in this period, and to use digital tools and methods to 'map' and visualize incidents of extreme weather: floods, storms, freezing conditions, and incidents including the 'year without a summer', in 1816 (after the Tambora volcano eruption caused severe climate abnormalities).

²⁰ Eira.llgc.org.uk. The partners in the project were the Centre for Advanced Welsh and Celtic Studies at the University of Wales, the ACRE Project at the UK Met Office, the department of performance at Aberystwyth University, and Welsh performance artist, Eddie Ladd.

The project also carried out community engagement, gathering interviews with local communities to share their experiences of extreme weather events within living memory: floods, storms, and snowfall. From these two strands of evidence, we were able to construct two types of interpretation and analysis: the scientific visualisations; and also a public performance piece, *Ghost Dance*, by Eddie Ladd: this drew on disparate narratives describing events of the winter of 1963 in an allegorical, not didactic way.

The project required sustaining a complex, hybrid archive of materials related to extreme weather impact in Wales, their analysis through performance and scientific visualization, and the communities that contribute this data. As such, it exemplifies the possibilities of digitally integrated research as set out in the Digital Humanities Manifesto: the ability to combine research, curation, and archive management to re-imagine the creation of digital content as a process of creative making, bringing together scholars and curators in ways that recast the role of each, driven by a shared goal of using historical sources for new purposes. These approaches effect collaboration of disciplines and data types as an act of curation as much as a piece of scholarship, one in which scholars will not just create not just a collection of sources, but also effect a convergence of practices: artistic, scientific, and humanistic, and the ability to work with connected communities around the content²¹. Documenting and publishing this sort of practice over the long term will also create a range of interesting issues – how do we replicate the relationships between archives, scientific visualization, and performance? How is provenance of archives retained when embedded in a scientific visualization, or a performance? How will the convergence of digital and born digital be preserved and replicated over the long term? This form of knowledge production is an exciting means of determining the interface to information, data, and knowledge. It also raises significant information management research challenges associated with managing and sustaining complex collections over time, pushing the boundaries of what is currently known about stewardship and curation of humanities and cultural heritage content and data.

²¹ <http://manifesto.humanities.ucla.edu/>.

Working in a culture of abundance: Europeana.eu

There are, therefore, a number of opportunities and challenges for researchers working in digital culture of abundance. And when it comes to abundance, Europeana.eu is the largest, and best known aggregator of European cultural heritage resources. So it may be that Europeana can offer some solutions to these identified challenges: helping researchers to select and understand content; supporting the creative use and re-use of content, and working with multimedia content in different disciplines and formats.

Europeana.eu is an online platform that acts as an interface to over 50 million items: books, paintings, films, museum objects and archival records that have been digitised throughout Europe, with the objective of making the 10% of digitised cultural heritage more usable for research, teaching, and public engagement, focusing especially on making it available for creative re-use releasing its impact for society and the economy. More than 2,000 institutions across Europe make their collections available through Europeana, including the Rijksmuseum in Amsterdam, the British Library and the Louvre to regional archives and local museums from every member of the European Union, including the National Library of Wales. In terms of full-text content, The Europeana Newspapers project is an interface to several million pages of text and images available in the public domain. Together, their assembled collections let users explore Europe's cultural and scientific heritage from prehistory to the modern day. Europeana has been aggregating metadata of digital heritage content, and making data (both content and metadata) openly available, for nearly 10 years. It recently produced a new portal for searching across this collection, offering new functionality and a greater degree of interactivity through services that are part of the 'platform', rather than a 'portal': Pro, Labs, and End User Services, making much of Europeana's dataset available for reuse via its API, portal and linked data sets.

For an example of how these tools can integrate content from disparate locations, WUD (What's Up Doc) is a customised search engine helping pull, link and organise data from two major cultural heritage repositories: Europeana.eu and the Digital Public Library of America (DPLA). It was developed as part of the research and outreach activities of the Medicine and Society chair at University of Fribourg. Initially intended for researchers: it stems from an effort to hand pick

images and documents for the *Medicine and food virtual exhibition*. Using online tools and APIS, the project team were able to streamline and make more efficient the curation process.

The Europeana Aggregation flow

Figure 3 shows the Europeana aggregation flow. The Europeana ecosystem depends on a network of national, thematic and domain aggregators which bring together, manage and provide access to data about Europe's cultural heritage. The concept of aggregation partners has been the cornerstone of Europeana's business model from the very beginning: Europeana does not currently have the resources to ingest metadata directly from the large number of organisations that already supply, or wish to supply, metadata to Europeana. The aggregator model makes it possible to obtain metadata from thousands of cultural heritage and scientific institutions while directly ingesting metadata from fewer than 150 organisations.



Fig. 3.

Europeana: Understanding researcher needs

The digitised content of Europe's galleries, museums, libraries and archives has huge potential for research, and Europeana has identified several ways it can do more to focus on researchers. From a technical and collections perspective, it is building focused aggregations of content, particularly full-text, and exploring how the content can

be re-used by the research community and highlighting collections in the Europeana dataset of specific interest to researchers based on some successful examples with Europeana Newspapers. Also ongoing is the exposure of Europeana's aggregations of text and metadata to allow research teams and infrastructures to build specific tools and services to fulfill particular research tasks, like looking for links between types of data, or doing the kind of nGram modeling described above. One of the biggest areas of potential is linking disparate collections and tools in different institutions and building workflows between them, although of course, issues of licensing, interoperability and access can often impede the re-use of that data in research.

Working with researcher communities of practice is also an important way of optimizing Europeana for scholarship, and understanding what specific communities need. EAGLE (the Europeana Network of Ancient Greek and Latin Epigraphy,) is an excellent example of this sort of liaison. EAGLE provides a single, user-friendly portal to the vast majority of the surviving inscriptions of the Greco-Roman World, as well as explanatory information and translations for the most important: a massive resource for researchers as well as the public. As part of Europeana, the EAGLE dataset includes over 1.5 million items, currently held across 25 EU countries as well as the east and south Mediterranean. EAGLE has supported the development of services such as a mobile application, allowing tourists to understand the inscriptions they find in situ by taking snapshots with smartphones, and a storytelling application that will allow teachers and experts to assemble epigraphy-based narratives.

A recent project to investigate optimizing Europeana Content for research, Europeana Cloud, conducted detailed analysis of the digital scholarship life cycle. Looking at a variety of user communities, the projects identified a number of key challenges to research use and potential of Europeana:²²

- Different research disciplines use different types of data in different ways, and there is no one size fits all solution
- Data aggregation should be aimed at horizontal rather than vertical solutions: rather than enabling deep down investigation of a topic, the access to aggregate data offered by Europeana is useful for comparison, sampling, and annotation of a variety of data from different locations.

²² <http://pro.europeana.eu/europeana-cloud>.

The value is in bringing together material, rather than exploring a complete collection in detail

- In order to encourage access to, and better discovery and knowability of collections, there is a need for enhanced collection-level descriptions
- There is a greater need for user-friendly tools and services which will enable re-use of Europeana data
- Digital humanities research requires access to greater amounts of data at the API level than is currently the case

In many respects, these reinforce the findings outlined above. The best way to address these in the Europeana ecosystem is developing a better understanding of how content and metadata is actually used, and their relationship with digital methods and tools. There is also a need for greater engagement with research communities, but ultimately, the least surprising finding is that there is a need for increased content of use to potential stakeholders. Once again, the requirement for more data is paramount.

Europeana Research

These findings have made development of Europeana Research a priority from 2016-8. Resources have been allocated to this initiative through the Europeana DSI funding strand to explore these issues, as it is seen as an important mechanism for connecting digital heritage with the research community. It has established a research Advisory Board to direct a series of funded activities to enhance the use of Europeana for Research, and to better document some of the innovation that can be achieved in re-using digital cultural heritage content, expressed in its mission statement:

“Europeana Research will help open up cultural heritage content for use in cutting-edge research. It will run a series of activities to enhance and increase the use of Europeana data for research, and develop the content, capacity and impact of Europeana, by fostering collaborations between Europeana and the cultural heritage and research sector. It will provide an important focus for the emerging communities of

practice who rely on Europeana for their research, and support the European investment in digital cultural heritage.”²³

The workplan for Europeana Research has several areas of focus, which will be overseen by the Europeana Research Advisory Board. There will be a *grants programme*, offering funds to researchers working with digitised content from the cultural sector that is part of the Europeana ecosystem, exploring how content can be re-used by the research community. The Europeana Research platform will highlight various *featured collections* available for re-use. These collections will be expanded and have extra descriptive metadata for discovery and use. There will be a focus on researching enhanced ways to describe *case studies* of successful use of Europeana content with DH tools and methods, developed through user engagement and dialogue with groups working in specific disciplinary areas.

Conclusion: Impact of Europeana Research on European Research Infrastructures

Europeana Research will work with research infrastructures such as CLARIN²⁴ and DARIAH²⁵ to help address strategic issues, and to encourage a flow of information on understanding engagement with digital content, tools, and methods. The grants programme especially will be a unique opportunity to engage with researchers directly and develop a better understanding of how they can work with digital content and metadata, developing a better awareness and understanding of available tools for analyzing data through the Europeana platform. This will provide an important layer of evidence that will help shape the development of research infrastructures in the humanities. Despite a significant, there is still concern that what Joris Van Zundert has called ‘big, all encompassing all serving digital infrastructures²⁶’ have little value for digital humanities technology development. He argues that what is needed for the humanities is flexible, small-scale research focused development practices, and highly specific distributed web services.

²³ <http://research.europeana.eu/>.

²⁴ clain.eu.

²⁵ daraih.eu.

²⁶ Zundert 2012.

In order to deliver the services that researchers in the humanities say they want: trusted, open content that is ‘pre-knowable’ from the perspective of quality control; the ability to liberate content from its mode of production and dissemination for reinterpretation and analysis by open tools; and working with multimedia content in an integrated way, liberating data from silos of content type, we need to try and find scalable solutions to defined problems. The specific ways that the development of Europeana Research can inform developments in Research Infrastructures are by providing the sort of ‘agile’ development mentioned above, especially testing the integration of data with tools and methods for analysis, through understanding the use of digital content at all stages in the research life-cycle, and in fostering better connections with GLAM organisations who provide access to their content. If research infrastructures are to have value for scholarship, there needs to be a clear role for scholars to contribute to a greater scholarly investigation and critique of the digital content life cycle, to create a more theoretical reflection concerning the role of the digital in humanities research. Europeana Research is an important opportunity for a praxis-based critical engagement that is the key to understanding the ways that the digital is affecting knowledge production. Hopefully lessons can be drawn from this programme of activities that will inform the future development of research infrastructures so they can support these agendas: drawing us into new collaborations, leading us to encounter new methods for engaging with content, and assisting in developing new insights into our cultural heritage.

Bibliography

- CORDELL, RYAN. 2015. “Reprinting, Circulation, and the Network Author in Antebellum Newspapers”. In: *American Literary History* 27.3. <http://ryan-cordell.org/research/reprinting-circulation-and-the-network-author-in-antebellum-newspapers/>
- ELL, PAUL S., AND HUGHES, LORNA M. 2013. “E-Infrastructure in the humanities”. *International Journal of Humanities and Arts Computing*, 7.1-2: 24-40.
- ENUMERATE. 2014. *Survey Report on Digitisation in European Cultural Heritage Institutions*. <http://www.enumerate.eu/fileadmin/ENUMERATE/documents/ENUMERATE-Digitisation-Survey-2014.pdf>

- HALL, GARY. 2012. "Has critical theory run out of time for data-driven scholarship?". In *Debates in the Digital Humanities*, edited by M. Gold, 128-132. Minneapolis: University of Minnesota Press.
- HITCHCOCK, TIM. 2013. "Confronting the digital: or how academic history writing lost the plot." *Cultural and Social History* 10.1: 9-23.
- HUGHES, LORNA M. 2012. "Live and kicking: the impact and sustainability of digital collections in the humanities". In *Proceedings of the Digital Humanities Congress 2012*, edited by C. Mills, M. Pidd. and E. Ward. Sheffield: HRI Online Publications, 2012. <http://www.hrionline.ac.uk/openbook/chapter/dhc-hughes> (accessed April 11th 2016).
- . "Finding the Belgian Refugees in cymru1914.org". 2016. IN *SOON GONE, Long Forgotten – uncovering British responses to Belgian refugees during the First World War*, special edition of *Immigrants and Minorities*, 34.2. The data for this graph is accessible at <http://dx.doi.org/10.5525/gla.researchdata.301>.
- HUGHES, LORNA M., PAUL S. ELL, GARETH A.G KNIGHT, AND MILENA DOBREVA. 2015. "Assessing and measuring impact of a digital collection in the humanities: an analysis of the SPHERE (Stormont Parliamentary Hansards: Embedded in Research and Education) project". *Digital Scholarship in the Humanities*, 30.2: 183-198. doi:10.1093/llc/fqt054.
- LIU, ALAN. 2012. "Where is cultural criticism in the Digital Humanities?." In *Debates in the Digital Humanities*, edited by M. Gold, 490-509. University of Minnesota Press, Minneapolis.
- NICHOLAS, DAVID, DAVID CLARK AND A. WATKINSON. 2014. "Log Usage Analysis: What it Discloses about Use, Information Seeking and Trustworthiness". *International Journal of Knowledge Content Development & Technology* 4.1: 23-37.
- NICHOLS, David and David Clark. 2015. "Finding Stuff." In *Is Digital Different*, edited by Michael Moss, Barbara Endicott-Popovsky and Marc Dupuis, 19-34. London: Facet, 2015.
- ROBEY, David. 2008. *Sustainability of AHRC-funded digital resources*. Available at: <http://www.ahrcict.rdg.ac.uk/activities/review/synthesis%20of%20sustainability%20issues.pdf>.
- . 2011. "Improving sustainability of publicly digital resources". In *Evaluating and the measuring the value, use and impact of digital collections*, edited by Lorna M. Hughes, 147-156. London: Facet.
- ROBINSON, Peter. 2014. "Digital Humanities: Is bigger better?." In *Advancing Digital Humanities: Research, Methods, Theories*, edited by Paul L. Arthur and Katherine Bode. Basingstoke: Palgrave.
- ROSENZWEIG, ROY. 2011. *Clio Wired: The Future of the Past in a Digital Age*. New York: Columbia University Press.
- VAN ZUNDERT, JORIS. 2012. "If you build it, will we come? Large scale digital infrastructures as a dead end for digital humanities", *Historical Social Research/Historische Sozialforschung* 37.3: 165-186.