Live and Kicking: The Impact and Sustainability of Digital Collections in the Humanities

by Lorna M. Hughes

Citation

Hughes, Lorna M. 'Live and Kicking: The Impact and Sustainability of Digital Collections in the Humanities'. In: Clare Mills, Michael Pidd and Esther Ward. *Proceedings of the Digital Humanities Congress 2012*. Studies in the Digital Humanities. Sheffield: HRI Online Publications, 2014.Available online at: <https://www.dhi.ac.uk/openbook/chapter/dhc-hughes>

Abstract

Digitisation initiatives in libraries, archives, museums and educational institutions have created a 'deluge' of data in the humanities that has transformed the information landscape and the way it is navigated for research and teaching. The use of digital collections for scholarship – using ICT based tools and methods – has been the basis of transformative and innovative research across the disciplines, allowing enhanced access to materials, and supporting new modes of collaboration and communication. Digital collections development has required a huge investment of public and private funding, and has, in turn, engendered an even greater requirement for continued investment to expand, enhance and augment the existing digital knowledge base, and to embed these collections in research and scholarship. More importantly, issues of sustainability and preservation of digital resources over the long term have an impact on the long tem use and impact of digital collections.

This paper discusses the research programme in digital collections at the National Library of Wales, and how it developing an evidence base for the use, value and impact of the digital collections of Wales, and investigating the relationship between impact and the long terms sustainability of digital collections.

Live and Kicking: The Impact and Sustainability of Digital Collections in the Humanities

by Lorna M. Hughes

1. Introduction

This paper discusses an initiative at the National Library of Wales to build a research programme around the digital collections of Wales. The programme seeks to document the impact and use of these digital collections for research, teaching and public engagement: and to develop new digital initiatives that will fulfil core remit of the National Library over the long term, set within to context of the "digital transformation of the Library" that "has caught everyone's attention" (McGann). The NLW research programme seeks to understand this transformation, and its impact internally and on the Library's users, and to address questions of sustainability of digital collections in the humanities. In order to frame this narrative, this paper will outline the development of technological innovation at the Library, and the ways in which digital collections and digital development are part of a continuum of the adoption of new technologies with a resultant effect on the mission of the organisation, and indeed on the institution itself. Understanding how digital collections are used is key to understanding their impact, and to making provision for the long term-sustainability of valuable digital content. Understanding the impact of the digital transformation in the Library also signposts how the institution itself remains relevant, with a continuing role as a custodian and advocate for documentary cultural heritage in an increasingly challenging economic climate – remaining not just alive, but kicking.

2. Technical Change and Cultural Narrative at the National Library of Wales

On Friday, June 7th, 1918, the Welsh weekly newspaper the *Llangollen Advertiser* carried a report of the half-yearly meeting of the Governors of the National Library of Wales, at Aberystwyth. Mr J. Herbert Lewis, MP, noted that it was gratifying that in spite of the adverse circumstances caused by
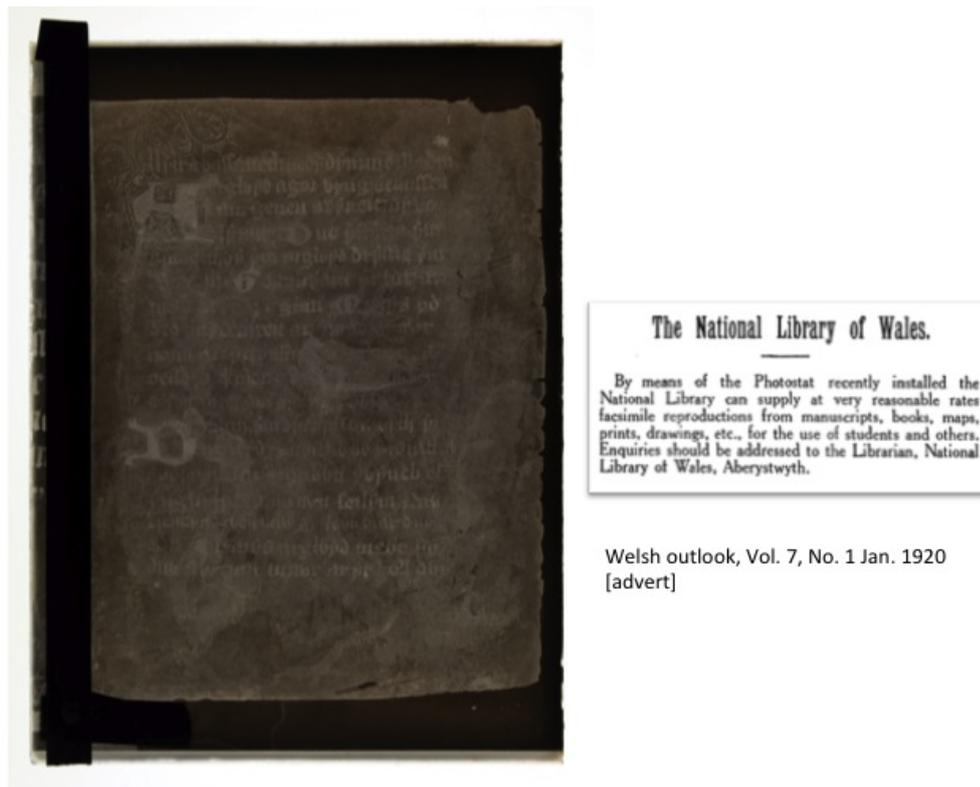
the war, the work of the Library had been successfully carried on under the direction of the Librarian, Mr. Ballinger. But the Library was doing much more than keeping calm and carrying on. Mr D.C. Roberts reported on behalf of the Finance Committee that the Library had received a pledge of funding of £500, from Major Lewis J. Mathias, of Aberystwyth, to purchase "a new photography equipment which would allow photographs to be taken of all manuscripts, and thus enable the schools in Wales to receive copies". This gift established the photostat and photographic section at the Library. In the same report, the Librarian referred to a collection the Library was developing of 'all records connected with the War so far as Wales was concerned', and 'appealed to the people of Wales to send all printed documents dealing with local efforts of any kind, as they would be invaluable when the history of the war came to be written'.[1]

In some regards, we can see in this report exemplars of two absolutely fundamental aspects of the National Library of Wales: an appreciation of the importance of state-of-the-art technologies for the delivery of the Library's mission, and an understanding of the importance of the creation, preservation and dissemination of uniquely Welsh collections, recognising the role of these collections in both disseminating and shaping the national cultural heritage narrative of Wales.

Figure 1: A glass negative of a photostat of a Peniarth manuscript, and the Library's advertisement for the photostatting machine.

The National Library of Wales.

By means of the Photostat recently installed the National Library can supply at very reasonable rates facsimile reproductions from manuscripts, books, maps, prints, drawings, etc., for the use of students and others. Enquiries should be addressed to the Librarian, National Library of Wales, Aberystwyth.

Welsh outlook, Vol. 7, No. 1 Jan. 1920 [advert]

1   Negative photostat print of NLW Peniarth 610 MS 191

1918 was an auspicious time for the Library to take on this new initiative. The Library had been established by Royal Charter as a Legal Deposit Library in 1907. During its brief history, it had seen controversy over its foundation – there was an inevitable struggle with Cardiff over its location, after which it was agreed that the Library should be located in Aberystwyth, on the west coast of Wales, mainly at the insistence of its primary benefactor Sir John Williams who wished to see it established in a Welsh-speaking part of the country. Construction of the Library building began in 1911. In the intervening years, Aberystwyth has not moved any closer to the rest of Wales, a land that is geographically divided between North and South, West and East by mountain ranges and rivers, a physical infrastructure no more conducive to connectivity among the regions than it was in Ballinger's day. A 2004 Welsh Tourist Board advertising campaign created by Wieden and Kennedy underlines this, promising the visitor to Wales "areas of outstandingly bad mobile reception". Wales is also a bilingual country, with strong Welsh language traditions in many areas, especially the north and west, yet almost entirely English speaking in the industrialised south.

The Library, has, therefore, since its inception, had to address access issues. Before taking up his role at NLW, Ballinger was a public librarian in Cardiff,

and he saw the benefits of extending library collections beyond the library walls – he was convinced that "libraries should not be ivory towers, but should, on the contrary, reach out to the community" (Jenkins). This objective can be seen in the development of a "book box" lending scheme to adult education classes in the industrial areas of south Wales (Baggs). During the First World War, this was extended to become a "books for the troops" initiative, which sent boxes of Welsh books to training camps in Wales and the border countries, but also to Welsh Army units in Macedonia, Egypt and Palestine (Jenkins).

The acquisition of the Photostat machine, and the ability to make and distribute copies of collections from the Library, was lauded as something that would assist greatly in the mission of the Library: it could support access to rare and fragile materials. It could aid education through the dissemination of materials to schools and what we would now call; "distance learners". It assisted collections management and unification: the Library began a process of obtaining photostats of material that would compliment or complete its materials through arrangements with other organisations, including the British Library, the National Archives, and private and public collections nationally and internationally. The newly-established Photostat and Photographic section became a central part of the Library. Academic journal articles published in the period 1920-1939 (available through Welsh Journals Online, *www.welshjournals.llgc.org.uk*) show the increasing dependence of scholars on photostats for their research. Interestingly, photostat use is documented and cited in these publications, and frequently used to illustrate articles. This documented use peaks in the period leading up the Second World War.

In 1931, technical innovation continued at NLW, as two fluorescence cabinets were acquired. Manuscripts which are difficult to read in ordinary light can generally be successfully read and even photographed while exposed to the ultra-violet rays of the fluorescence cabinet – this was notably used for research on the Hengwrt Chaucer carried out by Manly and Rickert up to the 1940s (NLW Wales Journal).

The next major technology to be adopted by the Library was microfilming, and the existence of the photographic section meant the Library was well placed to support this new approach. During the Second World War, the Library was the home of many of the treasures of the National Gallery and the British Museum, and other collections, which were moved to Aberystwyth for safekeeping, and stored in the Library's "Cave", a series of tunnels under the hill on which the Library stands, for the duration of the

War. In 1941, the Library obtained a microfilm camera and many of these visiting materials were microfilmed alongside Welsh content in NLW – partly as an insurance against the possible loss of the originals, but again, for access, preservation and as an attempt to better represent the information in collections to scholars. In history of these developments can be seen a clear focus on both acquiring Welsh content from elsewhere in photostat or microfilm form, with the objective of building totality in coverage of extant Welsh collections for scholarship for which NLW was a portal; and also the use of these formats to disseminate Welsh content nationally and internationally: microfilms of Welsh MSS were among those sent to the American Council of Learned Societies from 1941-1945. It was through use of these microfilms that many scholars came to know the iconic Welsh manuscripts of the Llanstephan, Cwrtmawr and Peniarth collections.[2]

3. Welsh Digital Collections

The use of digitisation technologies to increase and enhance access to the collections of Wales can be seen as a continuum of the enthusiasm and innovation attached to the adoption of new technologies – be they photostat or microfilm – throughout the history of the National Library of Wales, and as a pragmatic response to particular issues associated with the Library's mission, collections, history, and location.

The Library was an early and enthusiastic adopter of digitisation and networked technologies, and actively engaged in mass digitisation since 1998, creating free online access to digitised content from its collections. The scope of digitisation has included all the types of content in the Library. Print digitisation comes under the auspices of the "Theatre of Memory" programme, an ambitious initiative to digitise as much as possible of the Library's print collections. This includes *Welsh Journals Online/Cylchgronau Cymru*, the digitisation and online delivery of leading Welsh Journals in Welsh and English. This project was funded by JISC and the Welsh Government, who then went on to fund a more ambitious project, *Welsh Newspapers Online*, over one million pages of nineteenth century Welsh newspapers, again, in both Welsh and English. The Library has digitised highlights from its manuscript collections, and made available 180,000 page images of Welsh wills. Paintings, photographs and audio-visual materials have also been digitised and made available. As these source materials are varied and complex, their digital surrogates are increasingly multimedia in nature: text, image, moving image and audio. The National Library of Wales

has therefore made its own significant contribution to what has been described - in the context of scientific and other research (and business) data - as a "Data Deluge" (Hay and Trefethan).

4. Cyberinfrastructure in Wales

It is important to see this development of digital collections in Wales against the background of devolved government, which has, since 1999, sought to actively support the development of digital and other technologies in order to promote specific government policies intended to improve links between the geographically and linguistically disparate parts of Wales, and to create a shared sense of Welsh history, culture and nationhood. The Welsh Government initiative "Digital Wales"[3] has developed a framework for the integration of content, creativity and technology with an underlying drive towards the delivery of public services that are "digital by default" as a means of integrating services and people. To support this agenda, in 2010 the Welsh Government launched "Delivering Digital Inclusion: A Strategic Framework for Wales", which supports uptake of internet-based services and – interestingly - resources and content.[4] This has been described as "an attempt to connect communities and make available the culture of Wales: The Assembly is grasping the potential of the internet to facilitate some sort of 'public sphere' in Wales" (MacKay). This is an interesting analogy, building on the perspective developed by Habermas of a public sphere to be found in the institution of eighteenth century coffee houses of European capital cities (Habermas). A digital public sphere could be viewed as just as exclusive as coffee houses with a narrow customer base in terms of class and gender, however, setting our digital content in the context building a public sphere for Welsh communication and engagement is an interesting model.

This digital public sphere has also been supported by the development of highly collaborative frameworks, both human and technical, many supported by the Welsh Government division, CyMAL (Cymru Museums, Archives and Libraries) and the Welsh Higher Education Libraries Forum (WHELF), involving the librarians of all Higher Education organizations in Wales and NLW. WHELF has enabled a great deal of strategic co-operation, notably in the joint procurement of commercial digital resources, especially Journals and packages like Early English Books Online and British Library Newspapers (Prescott). NLW makes these available to all registered members with a Welsh postcode, creating what Andrew Green, librarian from 1998-2013 has called a "radical re-think of what it means to be a

Library member – building a truly national online knowledge service available to every Welsh citizen."

This collaborative framework also made possible the development of the Welsh Repository Network (WRN), funded by JISC and the Universities of Wales, which created an integrated network of repositories in all Welsh universities and the NLW allowing, including a service for storing and harvesting electronic theses (Prescott). The existence of the WRN means that a key layer of infrastructure for the sustainability of digital collections in Wales is already in place, subject to future development.

5. Digitisation of Welsh and Celtic Cultural Heritage

This framework for human and technical collaboration in Wales has been an important backdrop to the mass digitisation programmes of NLW. The Library took an early decision to use some of its core funding from the Welsh Government for digitisation projects, and this has enabled the Library to build internal expertise in the entire digital workflow, including all aspects of digital capture, management and preservation. This approach has now placed digitisation – and the need to sustain digital resources – as central to delivering its institutional mission in several key areas.

The most obvious is *access*, as digitisation offers broader access to the Library's collections to a global audience. As photostats and book boxes were once disseminated to audiences outside Aberystwyth, digital content takes our content worldwide. It also allows enhanced access, making collections searchable, findable and linked to related materials, in a variety of formats. Our audience comes from all academic disciplines, and their use of our content is both akin to the uses that are customary in the reading rooms (for example, the use of name-rich content by family historians) to the unforeseen (for example, the use of our photographic collections by media companies to produce posters and illustrations for many artifacts), and the mining of our newspaper collections by linguists tracking use of the Welsh language over time). As these communities expand, there is increasing demand for more collections to be digitised. Digitisation also supports collections enhancement through the 'virtual re-unification' of disparate content (not all the analogue copies of newspapers and journals in the digital collections are physically located in the library: the digital projects have enabled us to complete content with collections from elsewhere).

Digitisation also supports the Library's objective of *preservation*, by

providing digital surrogates of rare and fragile materials, including manuscripts. Expertise in managing the entire digital life-cycle has also enabled the Library to develop a better understanding of the collection of born-digital materials – many of the personal, artistic, and literary archives the Library has acquired recently consist of laptops, obsolete computer disks, and historic software. For example, the Brith Gof performance archive[5] has come to the Library in the form of written papers, drawings, architectural models, as well as computers, LP records, and moving image material on a variety of digital video and video cassette formats. Digitisation of this material for wider access is also a process of 'digital archaeology', and a process of ongoing investigation of formats and understanding how continued access to complex material can be fostered. Digital development also supports *collections development*: by building capacity in this area, the Library is able to continue to acquire new and more complex materials, such as the Phillip Jones Griffiths archive, and the archive of ITV Wales (the Welsh commercial broadcaster).

Figure 2: Some items from the Brith Gof archive.

Most significantly, digitisation is creating content that can *transform scholarship* across the disciplines through the use of digital tools and methods for the analysis and re-use of this content for research and education (Hughes). This has been the driver for the most ambitious of the NLW digital projects: the digitisation of Welsh Journals and newspapers; a current project funded by JISC to digitise the Welsh Archives of the First World War; and the ultimate objective of digitising the Library's entire manuscript collection.

Underpinning all NLW digitisation is the underlying principle of freely available public digital collections. The free provision of all digital content is a firm commitment by NLW, in keeping with the goal of free Library provision. This programme of digitisation and delivering information online, in the context of a national drive for digital literacy and engagement, can be seen a move towards delivering core Library services as "Digital by Default".

Given this significant investment in digital collections there is increasingly a need to demonstrate their use, impact and value, especially set against the increasing pressure to charge for digital content due to the vast investment in digitisation by commercial publishers. There is a need to show that digital collections represent good use of public funds, that they facilitate engagement with new user communities, and that they are the basis for innovative use of digital content in research and teaching. Similarly, a better understanding of the use of digital content can help the custodians of digital content plan for the sustainability and long term preservation of these valuable cultural artefacts. This understanding is even more important in the current economic climate of unprecedented and devastating cuts to cultural heritage and academic programmes, including those building and delivering digital content and services. This is the case at the National Library of Wales, which has seen significant budget cuts each year since 2010.[6]

Despite the importance of the need to gather evidence of the impact and use of digital content, currently evidence is sparse. The exception is systematic work undertaken recently at King's College, London (Deegan and Tanner) to assess the impact of digital collections using a tool called a "balanced value model".

6. Understanding the Impact of the NLW Digital Collections

The research programme in digital collections at NLW has been exploring these issues, and carrying out a series of projects to investigate and understand the use and users of digital collections, and to feed these findings into the development of new digital content. There have been considerable challenges in developing this work, not least, the reliance of digitisation initiatives on short term funding. Digitisation projects are generally funded over a short period of time – two yeas in the case of *Welsh Journals Online*, 18 months in the case of the *Welsh Experience of The First World War.* At the end of the funded period, the project is launched, staff move on, and it is difficult to carry out a retrospective understanding of the impact of the resource. This short-termism is at odds with the reflection required to understand impact, which accrues over time, and requires resources that are seldom available after a project is launched. The recent ITHAKA report, *Sustaining Our Digital Future*[7] showed that after the launch of a digital resource, little attention is given to issues of post-grant sustainability that extends beyond data management. There is a perception that revisiting a digital resource to understand its impact, assess its user base, and enhance its embedding in teaching and research, are luxuries that institutions can rarely afford. This lack of systematic research on the impact of digital collections on scholarly practice contrasts strongly with not only the investment in digitisation, but in funding of research into e-infrastructures, e-science, and the development, management, funding and sustainability of digital repositories (Nicholas et al).

In 2011, a report on the impact of digital resources on scholarly information seeking in the humanities was created for the UK's Research Information Network (RIN). This report suggested that humanities scholars are enthusiastic users of digital resources, but they are highly selective and demanding of digital content, and easily deterred from using digital resources: in order to be useful for research, these must be of a high quality, useable, and with easily identifiable provenance. These factors: quality, usability, and provenance are all qualities inherent in digital content that can be understood and documented using a combination of qualitative and quantitative methods. A method for bringing these approaches together is the Toolkit for Digital Scholarly Resources (TIDSR), developed in 2009 by the Oxford Internet Institute (OII) through funding from JISC to develop an impact assessment mechanism for the first phase of JISC digitisation

projects. TIDSR gives access to a set of open approaches and tools for Quantitative methods - using analytical tools (like Google Analytics, or deep log analysis) that provide statistics about the 'footprint' of a resource, including traffic and visitors to a web resource. – and qualitative methods, such as interviews and case studies, to provide more fulsome answers. TIDSR is an integrated approach to measure and understand all factors that affect 'use', enabling us to not just value things we can measure, but to understand how to measure things that we value. It supports the measurement and evaluation that affect impact: types of use; awareness of the resource; frequency of citation; the effectiveness of marketing strategies; and the embedding of the resource in research and education. Some of these can be counted, using quantitative methods, including Webometrics, Bibliometrics or Scientometrics; Log file analysis; referrer analysis, and content analysis. More detailed answers and reflective input can be obtained through qualitative methods, including Focus Groups; Interviews; and User Feedback. The findings are then brought together in a detailed 'impact assessment' that 'triangulates' findings from the different methodologies, gathering data from the perspective of different user communities (Meyer et al).

In 2010, the JISC e-Content programme's "impact and embedding" scheme funded a team from King's College London, Queen's University Belfast, and the University of Strathclyde to use TIDSR methods to carry out an impact assessment of the *Stormont Papers* resource, an AHRC funded collection of the devolved Northern Irish Parliamentary papers from 1921-71. The SPHERE (*Stormont* Parliamentary Hansards: Embedded in Research and Education)[8] project was an opportunity to assess the effectiveness and replicability of the toolkit for measuring impact of a humanities resource. The quantitative methods and numeric data did not have a steep learning curve, even for a team that lacked a rigorous training in statistical methods, especially when combined with the stakeholder engagement and other qualitative approaches (for example, interviews explained the reasons that visitors from certain communities used the resource, and their patterns of searching).

The most instructive finding of the SPHERE project was that the TIDSR toolkit was a simple to use and easily replicated methodology for understanding impact of a digital resource. The Research Programme in Digital Collections at NLW is now using it for impact assessment of our digital collections, and some preliminary analysis has now been carried out on the *Welsh Journals Online*. A small study focused on analysis of use of the resource over a six-month period from January-July, 2012. The resource was

launched in 2009, and linked to the JISC teaching resource JORUM in late 2011 *(www.jorum.ac.uk/),* so this timeframe is a suitable period to gather a snapshot of use based on the weblog data. The study addressed the popularity of the resource, in terms of:

- Quantifying the number of new and returning users that access it.
- Identifying the user communities that utilise the resource.
- Documenting the route that users take to locate and access the Welsh Journals web site.
- Finding evidence of the use of *Welsh Journals Online* in academic research papers.

Statistics for usage of the resource (Table 1) were relatively high. Visitor loyalty was also encouraging – most visitors viewed a number of pages:

*Table 1: Statistics for usage of the Welsh Journals Online.*

	Welsh	*English*
Visitors		
Total Visitors	39,528	79,070
Average Visitors per Day	140	281
Total Unique IPs	16,962	41,101
Page Views		
Total Page Views	2,131,686	4,558,273
Average Page Views per Day	7,586	16,221
Average Page Views per Visitor	53.93	57.65
Hits		
Total Hits	4,296,935	7,757,327
Visitor Hits	2,380,482	5,390,500
Spider Hits	1,916,453	2,366,827
Average Hits per Day	15,291	27,606
Average Hits per Visitor	60.22	68.17
Cached Requests	30,436	127,458

Analysis of the countries that users came from showed that visitors came

from mostly the UK (61%).

Referral data showed that for both the English (Figure 3) and Welsh (Figure 4) language version of the site, users mostly came from websites that were resources for genealogy. This is unsurprising, given the fact that family historians make up a large proportion of the NLW user community, and also given the popularity of the journal titles for family and local history. Nonetheless, it indicates that the resource should be better promoted with other communities, especially academic users.

Figure 3: Top referral sites for Welsh Journals Online (English site).
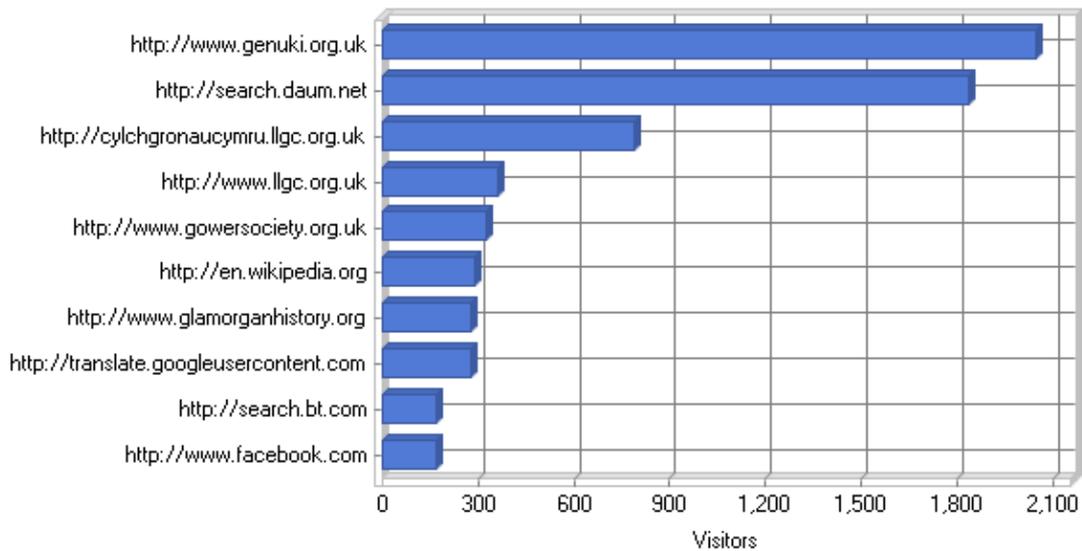
Figure 4: Top referral sites for Cylchgronau Cymru (Welsh site).
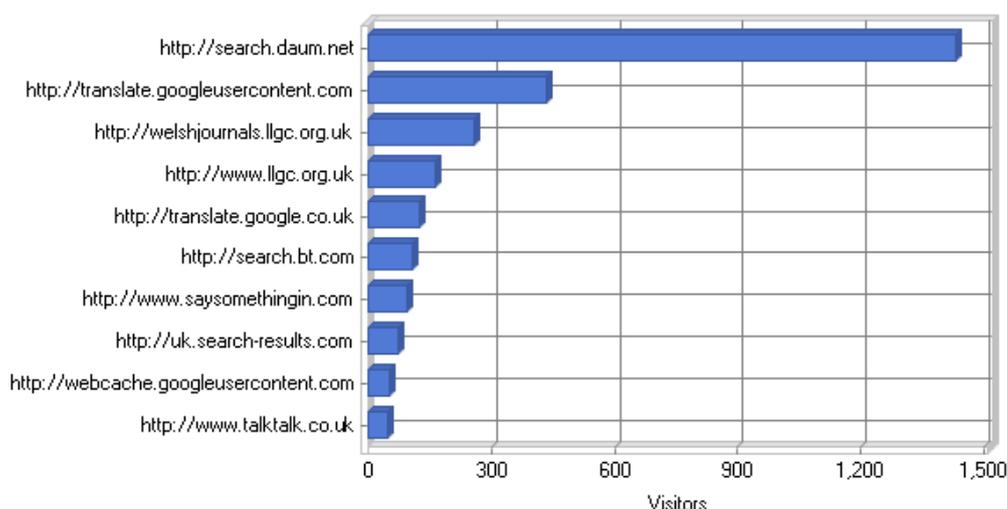
A small experiment with Content Analysis methods was also carried out, using the TIDSR methods for citation analysis, using online sources of bibliometric data to track citations to a particular resource. This method can be used to locate textual content that referred to *Welsh Journals online*, to understand the extent to which the resource is reported to the wider user community. This was primarily intended to investigate citation of the resource. Google Scholar found only one explicit citation of the digital resource since it was launched in 2009. However, there were sixty-two citations that referenced the individual English medium journals held within the resource during this period. A series of interviews with a selection of authors who could be contacted for preliminary investigation expanded this finding: most of them had indeed used the print journals, as the research had been undertaken some time before *Welsh Journals Online* had been released. However, a small number of authors had indeed used the online resource for the reference, but cited the paper journal, mostly because they "did not know how to cite the electronic resource". Feedback indicated that they would welcome the addition of a tool or mechanism to enable citation. However, citation analysis in the context of journals is subject to many vagrancies, given the long lead time for journal articles to appear, so these findings should be regarded as preliminary only, to be revisited once the resource is more mature and patterns of citation can emerge over a more statistically meaningful period of time. If we look back at citations in the journal references to Photostats, these peaked about twenty years after the service was introduced, in the late 1930s. There is a long lead-time for new resources to make their way into scholarly publishing – and again, this is an issue that makes analysis of 'impact' of the digital humanities collections difficult – true impact emerges over time. Value, impact and use of

collections take time to develop, and this needs to be reconciled in a world of responsive, short-term funding opportunities.

While the findings of this analysis are presented here in a summary, they do illustrate that this sort of impact assessment exercise is useful to do as a post-project "health check", to get a better understanding of the use and users of digital collections. This type of exercise should be applied to all digital resources post-launch, but the reality is that this is rarely carried out. Partly, this is due to a lack of post-project resourcing, but also a lack of awareness of the need to collect user data and log statistics. In the review of the sustainability of AHRC funded projects carried out by the AHRC ICT Programme in 2008 (Robey), only half of the projects funded planned to collect user statistics at the end of their project.

This is problematic, as there is an important relationship between use and long term sustainability of digital resources: *Sustaining our Digital Future* shows that digital resources that are used are more likely to be sustained over the long term. In order to develop a better perspective on how digital collections fit into the scholarly life cycle, we need to understand the use and users of digital collections. It may be that we have placed too much emphasis on technical sustainability, the provision of research infrastructures and technical management of the digital life cycle as topics for research, and that more investigation of the use of digital content in scholarship is necessary in order to safeguard our valuable digital assets. To investigate this theory, ITHAKA, in collaboration with NLW developed a tool to check the health of digital projects post-launch that encourages the custodians of these projects to take stock of use of their resources at regular intervals after they are launched.

However, the main finding from these studies using TIDSR to understand impact of *The Stormont Papers* and the *Welsh Journals Online* is that carrying out impact assessment on an existing digital resource is frequently too little, too late. Impact can be measured retrospectively, but should be planned at the outset of the development of a digital resource. The impact of digital collections will be increased if out a thorough analysis of the factors that will enhance impact is carried out at the outset of a project – ideally, before seeking funding. This will identify crucial aspects of a project that will increase its impact, whether this relates to the content selected for digitisation, approaches to target specific user groups, or interface enhancements to assist usability. If impact is *only* assessed once a project has ended, there are usually few resources available to implement emerging recommendations.

7. Embedding Impact into Digital Project Development at NLW

The Welsh Experience of World War One/Rhyfel Byd 1914-1918 a'r profiad Cymreig

At NLW, the findings of the two studies described above and analysis of existing research in this area (see Hughes), are providing an empirical basis for the development of a new mass digitisation project funded by JISC, The *Welsh Experience of World War One/Rhyfel Byd 1914-1918 a'r profiad Cymreig*. This project, a collaboration with the special collections and archives of Wales, will digitise 190,000 pages of previously 'hidden' primary source materials related to the impact of WW1 on all aspects of Welsh life and culture from 1914-18.[9]

In developing the project, we identified specific actions to enhance impact throughout the digital life cycle.

The content selected for digitisation was indentified through extensive pre-funding engagement with academic research specialists on the First World War, who advised the project team on the archival source materials that should be selected for digitisation as the most useful for research and education. The project has also convened a multidisciplinary 'interface development group', comprised of academics, as well as archivists, librarians, (including custodians of the original analogue collections) digital humanities specialists, developers, and education specialists. This group has provided a range of perspectives on interface development of the resource to ensure that it is effective, serves users' needs, and will support resource discovery as well as embedding the resource into research and education. In order to accomplish this, the focus will be on building something that is clear and simple, yet flexible, designed to allow serendipitous discovery of materials by users, reflecting the experience of working with an analogue archive, and for flexible re-use of the content for as-yet unforeseen uses. This will also support the integrity of the resource, by retaining the narrative of the archival selection, making it clear to users what has been selected and why.

Another significant aspect of the project is that it is based on ingest of the digital content into an underlying technical repository, making use of the NLW's FEDORA-based digital repository architecture. This supports the archiving of multimedia digital content; and the exposure of content for harvesting and aggregation. Existing NLW workflows have been modified to

allow ingest into the repository of new content types all content types created by the project: printed text; newspapers; photographs; manuscripts; audio, and moving image materials. Once ingested, digital resources are given permanent and persistent identifiers and are subject to preservation management actions. Metadata for the digitised resources is exposed for harvesting and aggregation using the OAI Protocol for metadata harvesting, and an API layer will also make the content accessible.

This anticipates strategies for sustainability at the outset: Building an open platform as the infrastructure for the project will support the integration of additional content if needed (should future funding for digitisation should become available).

The resource will also be linked to from key external digital resources that relate to the First World War, in collaboration with national and international WW1 Commemoration activities (including the Strategic Content Alliance and Welsh Government WW1 commemoration initiatives)[10] as well as international portals and content aggregators including Europeana (www.europeana.eu). Inclusion within trusted gateways to similar content will increase uptake and use – and therefore impact – of the resource.

Practical methods for measuring impact have also been incorporated to the project from the outset. For example, statistics about the usage of the original analogue materials have been gathered, for a point of comparison with their digital surrogates. Again, scoping the content in collaboration with academics has helped identify digital content that will fill specific gaps in our knowledge and understanding of the topic and foster new areas of scholarship, so it should be easier to monitor subsequent embedding in research and teaching. These measures – as well as ensuring that web statistics and log analysis data are maintained – should provide the core data required to assess the impact of the resource several years after its launch. This data will ideally enable us to enhance the resource where necessary to continue to update and modify the content and functionality, and to increase its use by diverse groups of users.

8. Conclusion

The creators and custodians of digital resources need to put use and users of digital collections at the forefront. If digitisation is to have more impact than being a form of "digital photocopying" (or even digital Photostatting), the

user needs to be placed at the centre of the process from the outset. In the humanities, if we are creating resources for scholarly, academic use, then the research community must have input into the selection of content for digitization; the use of this content for research and teaching; and how it is to be discovered. This will increase the impact and value of digital collections, and therefore, the evidence increasingly suggests, increase the likelihood of their sustainability. We also need to be clear about what we mean by impact: scholarly digital resources created by the sort of funding available to the humanities cannot compete with the scale of resources created by commercial publishers, and it is unrealistic of funders to expect that they will. In comparison to this material, what we can produce will be inevitably be seen as niche, so we must maximise its impact and value through careful selection and targeting, ensuring that it contains the material that will most serve the community for which it is intended, or allow hidden sources to be uncovered, or advance knowledge and education by linking and providing access to materials in new and unique ways. Good economic thinking about the sustainability of scholarly resources is not limited to the means of financing, but to the value of the use of the resource by current and future users. This is a perspective that will enable us to view the creation, provision and sustainability of digitised knowledge as a public good (Waters).

NLW has a long history of user engagement. Sir John Ballinger knew that his book box plan would be valuable, because he had run public libraries in south Wales and knew there was an audience for this service. The value of Photostat and ultra-violet photography can be seen by its references in journals, indicating the suitability of these technologies to the needs of scholars. The data from the analysis of the Welsh Journals shows that they are reaching the existing constituency of family historians. The NLW research programme in digital collections is now able to undertake engagement with scholars to identify new materials for digitisation, and build scholarly use into the organization and governance of our digital collections. We need to foster use and uptake of digital collections, and embed their use into scholarship, and to think of these not as ancillary activities, but an essential part of delivering digital collections. Champions for the use of digital collections will be as central to their sustainability as technical experts and cyberinfrastructures.

9. References

Baggs, C. M. *'The National Library of Wales book box scheme and the South Wales Coalfield'*, National Library of Wales Journal, 30, 1997.

Editorial Notes, National Library of Wales Journal, Volume 1 Summer 1940, issue 3
http://welshjournals.llgc.org.uk/browse/viewpage/llgc-id:1277425/llgc-id:1277591/llgc-id:1277601/get650/photostat%20chaucer

Green, A. "The Future of National Libraries and Archives". In Osmond, J. (ed.) Myths, Memories and Futures: the National Library and National Museum in the story of Wales, www.llgc.org.uk/fileadmin/documents/pdf/future_national_libraries_archives.pdf. 2007.

Hughes, L.M., Ell, P., Dobreva, M., and Knight, Gareth, K. *"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"*, Literary and Linguistic Computing. Oxford, 2013.

McGann, J. "Online Humanities Scholarship: The Shape of Things to Come", Houston: Rice University Press, 2010.

Meyer, E. et al. "Usage and Impact Study of JISC funded Phase 1 Digitisation Projects & the Toolkit for the Impact of Digitised Scholarly Resources". University of Oxford: Oxford Internet Institute, 2009.

Habermas, J. *The Structural Transformation of the Public Sphere*. Trans. Thomas Burger. Cambridge, Mass: MIT Press, 1991.

Jenkins, D. A Refuge in Peace and War *: The National Library of Wales to 1952*. National Library of Wales Press, 1997.

Johnson, Andrea.) *A National Audience Development Plan for Wales*. Andrea. Johnson Associates, 2008.

Hey, T. and Trefethen, A. *The Data Deluge: An e-Science Perspective*. In Berman, F., Fox, G. and Hey, T. (eds) Grid-Computing – Making the Global Infrastructure a Reality, Wiley, 809. 2004.

McKay, H., 'The Internet and Wales: Identity, Politics and Language'. Paper presented at Internet Identities and Europe conference, University of

Sheffield, September 2004.

Prescott, A.J. *The Digital Library*, in "Evaluating and measuring the value, use and impact of digital collections", ed. Hughes, L.M., Facet, London, 2011.

Robey, D. Improving sustainability of publicly digital resources, in LM Hughes (ed), *Evaluating and the measuring the value, use and impact of digital collections*, Facet, London, 2011.

Waters, D.J. Building on success, forging new ground: The question of sustainability First Monday, Vol 9, 5-3, May 2004. http://www.firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/1148/1068

Footnotes

[1] "National Library of Wales, Llangollen Journal and Merionethshire Journal, No. 2986, Friday June 17th, 1918. I am grateful to Nia Williams, digitisation assistant at the National Library of Wales, for spotting this reference during the process of newspaper digitisation for the JISC-funded "Welsh Experience of World War One- Rhyfel Byd Cyntarf 1914-1918 a'r profiad Cymreig" (cymruww1.llgc.org.uk)

[2] http://www.llgc.org.uk/manuscripts.

[3] http://wales.gov.uk/topics/businessandeconomy/digitalwales/?lang=ens.

[4] http://www.digitalinclusionwales.org.uk/

[5] http://humanitieslab.stanford.edu/BrithGof/Home

[6] See the NLW statement on the impact of budget cuts, 13th January 2012: http://tinyurl.com/czxnc44

[7] http://sca.jiscinvolve.org/wp/files/2013/01/Sustaining-our-digital-future-FINAL-31.pdf

[8] See http://microsites.oii.ox.ac.uk/tidsr/case-study/353/stormont-parliamentary-hansards

[9] The project is led by NLW, in partnership with archives and special collections of Bangor University; Cardiff University; Aberystwyth University; Swansea University; University of Wales Trinity Saint David's; The People's Collection Wales; Archives and Records Council, Wales (ARCW); and BBC Cymru Wales. It has received 500,000 in funding from the Mass Digitisation strand JISC e-Content programme (2011-13), and will be completed in October 2013. For more information, see: http://cymruww1.llgc.org.uk/

[10] http://wales.gov.uk/newsroom/firstminister/2012/120808ww1/?lang=en