CENDARI's Grand Challenges: Building, Contextualising and Sustaining a New Knowledge Infrastructure

Jennifer Edmond, Trinity College Dublin

Autobiographical Note:

Jennifer Edmond if the Director of Strategic Projects in the Faculty of Arts Humanities and Social Sciences at Trinity College Dublin. She coordinates the European Research Infrastructure Project CENDARI (2012-2016) and the Scholarly Publishing Working Group of the ESF Network on Digital Methods for Arts and Humanities (2011-2015).

Abstract:

In its widest sense, infrastructure allows us as finite individuals to achieve beyond our individual capacity to know, to do, to see.   But even within the more narrow context of research infrastructures, broad and diverse definitions exist of how such an infrastructure should deliver these enhancements in knowledge and perspective. The Collaborative EuropeaN Digital Archival Research Infrastructure (CENDARI) project was launched in 2012 to address some of the gaps in provision for digital historical research, building on a long tradition of work in libraries, archives, digital humanities research centres and other research infrastructures.   What distinguishes CENDARI, however, is its focus on what the project team has identified as the 'grand challenges' for each of its contributing stakeholder groups: collections experts based in libraries and archives, historians of the medieval and modern periods, and e-Scientists. This ethos, combined with its close relationship at European level to DARIAH, the Digital Research Infrastructure for the Arts and Humanities, has instigated a unique and fruitful approach to supporting historical research with digital resources, tools and spaces.

CENDARI's Grand Challenges: Building, Contextualising and Sustaining a New Knowledge Infrastructure

In its widest sense, infrastructure allows us as finite individuals to achieve beyond our individual capacity to know, to do, to see.   Although the term 'infrastructure' may be a twentieth century invention, the function of allowing valuable resources to be leveraged widely and publicly is of course much older. Roads are the classic example, where a cart track shared by generations is improved over time and eventually taken on as a public good. But physical infrastructures, with their visibility, tangibility and omnipresence, are of course not the only categories of shared investment in widely accessible capacity-raising, and some of these more subtle examples also have a long history. While the term 'soft infrastructure' has been coined to differentiate the tangible physical 'hard' systems from social systems, such as government, education and law, 'knowledge infrastructures' are also a part of our daily lives, and have been since villages supported the travelling bards as they visited to share stories, and since newspapers became an inexpensive way for the individual to understand their place within the experiences of their community and country.

The Collaborative EuropeaN Digital Archival Research Infrastructure (CENDARI) is also a project within this lineage of knowledge infrastructure. But before the project itself can be usefully discussed, it is important first to present some of the historical trends and current environment which contributed to its shape and vision. A useful starting point is with libraries and archives, classic examples of a point where knowledge infrastructures branch off into a smaller subgroup, that is, research infrastructures: a place where knowledge is available across a wide range of media and subjects, and where, in contrast to the bard and the newspaper, users can come with their own particular questions and hope to find answers they themselves, rather than the storyteller or editor, define as satisfying. Indeed, as the 2011 ESF report on infrastructure for the digital humanities makes clear: 'it was in the field of Humanities that the idea of an RI was first born.' [1] Although 'research infrastructure' as a category now itself has 'harder' and 'softer' manifestations with a myriad of vying definitions, the library remains, for the humanities researcher, the ultimate (and original) place for research and investigation. It is the cauldron from which the changes in the disciplines

emerge, changes which can manifest themselves as new methods or even new areas of study, which colour how and what scholars read, and shapes how they exchange ideas with colleagues and students. One might be inclined to assign the library or archive, with its imposing buildings and its shelves of book objects, to the hard side of the infrastructural continuum, but even the briefest investigation into the trends of the modern library, with its virtual collections and electronic journals accessible from around the world, and its physical spaces becoming ever more dedicated not to book storage but to spaces where users can engage and co-create knowledge, exposes the truth that our libraries and archives are valuable not for the physical objects they hold, but the information and potential knowledge those objects represent.

The road from physical library or archive to its virtual incarnation, from the humanities infrastructure to the humanities cyberinfrastructure, has not been a smooth one, however. Indeed, it has been suggested that the 'long now' of cyberinfrastructure:

> …is about 200 years. This is when two suites of changes began to occur in the organization of knowledge and the academy which have accompanied – slowly – the rise of an information infrastructure to support them: an exponential increase in information gathering activities by the state (statistics) and knowledge workers (the encyclopedists) on the one hand and the accompanying development of technologies and organizational practices to sort, sift and store information.[2]

While the most apparent modern aspects of the shift from written text to electronic data set has been a clear one in some fields, obviating the needs for phrases such as 'digital astrophysics,' we still find ourselves referring to 'digital humanities,' and experience the concomitant struggle between the old model and the new. For the libraries and archives still exist, and must continue to do so: the analogue record will never be as complete as a digital facsimile, and the threat that we might lose aspects of that record, from the material characteristics of the ink and paper to ways of linking objects that are other than linear, once material is digitised is clear.   What is difficult, however, is making a clear argument for integrating not the old with the new, but the new with the old. Research libraries have long been the lasting homes and preservers of research outputs, such as scholarly monographs and theses. But when scholarly outputs begin to appear ever-increasingly in the form of data sets, annotated virtual collections and web-based resources, the capacity (both in terms of skills and in terms of funding) of the library to maintain this output is challenged. The unknown cost of future maintenance and technical migrations is one of the unforeseen

challenges that came with the rise of the digital humanities, in stark contrast to the data-intensive sciences where long term preservation is less of an issue, as data is seen as much more transient and as having a far more limited span of utility.

This recognition of a shifting landscape for the underpinning of research and its outputs arrived alongside a number of macro-level changes in the European political arena. But these changes also exposed one of the most longstanding sources of tension in the European Research Area: how do we avoid investing only into the more easily identifiable needs of science, technology, engineering and medicine, the so-called 'STEM' subjects, and thereby further disadvantaging scholarship into the all important questions of who we are, where we have come from, and how we express innately human drives and capacities: that is, the arts and humanities? The clear intention to plan for a longer term, and to coordinate investment in research infrastructure focused minds on this question, and on the validity of the oft-heard statement that the arts and humanities simply did not need modern infrastructures. Going back twenty years, to the time when the dominant paradigms for many of the humanistic fields could be loosely grouped under the umbrella term of 'critical theory,' the question of what an arts and humanities research infrastructure could look like would have led to an interesting but unsatisfactory answer: as the various methodological approaches were so strongly associated with the thinking of specific individuals, or indeed with specific texts from their oeuvre, one could have imagined a system setting up the leading minds in an institute for all to access, where scholars could go and thrash out ideas in real time. But this model sounds strikingly familiar, indeed, very much like a university, that other classic knowledge infrastructure.

It was the rise of new methodologies, driven by the potential for discoveries within the humanistic domains through the applications of technology, which in the end provided a satisfactory focus to the question of what an arts and humanities research infrastructure could provide that libraries, archives and universities did not already cover. Canadian scholar Geoffrey Rockwell dates this movement, which he calls the 'infrastructure turn'[3] to a point in time somewhere between the establishment of the Canada Foundation for Innovation in 1997 and the publication of the American report *Revolutionizing Science and Engineering Through Cyberinfrastructure*, commonly referred to as the 'Atkins Report.'[4] In an European and a

Humanities context, however, the critical moment is more usefully dated to 2006, with the publication in 2006 of two landmark documents about research infrastructure: the first European Strategy Forum on Research Infrastructures (ESFRI) *Roadmap*[5] and the American Council of Learned Societies (ACLS) Report on Cyberinfrastructure for the Humanities and Social Sciences, *Our Cultural Commonwealth.*[6]

The first of these is remarkable both for its ambition, and for its hesitation. The ESFRI mandate was an enormous one: 'to describe the scientific needs for Research Infrastructures for the next 10-20 years, on the basis of a methodology recognised by all stakeholders, and take into account input from relevant inter-governmental research organisations as well as the industrial community.'[7] And, six years on from it, one has a sense when rereading that report of the burden of this mandate: future revisions and reviews of the Roadmap are already described against a backdrop of concern that European centres of excellence are too small, too fragmented, and too dependent on national infrastructures for European-level science. But the potential in centralised planning for coordinated development of shared, large-scale research infrastructures is also apparent, with its foreseen ability to increase the excellence, competitiveness, policy relevance and capacity of European scientific outcomes.

Where the ESFRI roadmap still seems tentative in 2006, the ACLS report on cyberinfrastructure for the humanities is strident, defining five clear aspects of what a humanities cyberinfrastructure must have -- from accessibility and sustainability to the ability to support collaboration and experimentation -- and putting forward eight concrete actions for the community to undertake in order to deliver on the priority of cyberinfrastructure for the humanities. But the clear recognition that dedicated cyberinfrastructure to support contemporary humanistic scholarship was required did not answer the question of what the shared support structure should look like. In the period immediately preceding and following the 'moment' of 2006 discussed above, a proliferation of projects emerged, each representing a view on what the platonic ideal of a digital humanities infrastructure should be. Suites of text analysis tools, such as HyperPo or TaporWare, had been in existence since the late 1990s, but whether these could or would be viewed as an infrastructure, or a public outgrowth of a limited project, is unclear. The idea of facilitating not just a known set of interventions,

but the whole of research in this crowded field of projects and visions, and to create

sustainable opportunities for enhancing the research environment, was too large a task for

many project leaders to feel comfortable with. In other cases, the ambition either grew out of

or preceded the developments. Although founded in 2002, the NINES (Nineteenth Century

Scholarship On-line)[8] project, now arguably an invaluable central resource for scholars of the

era it serves, was still defining itself in 2006, as it was only in 2005 that their key contribution,

providing peer-review for digital work in the (long) nineteenth century, emerged. Similarly, the

Mellon-funded Bamboo Project,[9] only began its initial scoping workshops in 2008.

One of the greatest challenges for these many projects on the threshold between

being a tool or discrete project and becoming an infrastructure is the ability to interlock with

existing practices 'below the level of work,' for it is only at this level that they can become both

omnipresent in scholarly work and simultaneously transparent. To be present at this level

means to operate:

> without specifying exactly how work is to be done or exactly how information is to be
> processed (Forster and King, 1995). Most systems that attempt to force conformity to a
> particular conception of a work process (e.g. Lotus Notes) have failed to achieve
> infrastructural status because they violate this principle (Grudin, 1989; Vandenbosch
> and Ginzberg, 1996). By contrast, email has become fully infrastructural because it can
> be used for virtually any work task. [10]

Beginning from a project-informed stance made it difficult for the infrastructural leap to occur,

and the essential requirement for collaboration between experts in the technical possibilities

and in the humanistic requirements was, and remains, one of the most difficult 'scaling

factors' in humanities cyberinfrastructure.

It was in this fluidic state that DARIAH, the Digital Research Infrastructure for the Arts

and Humanities, named in the ESFRI roadmap, was preparing and initiating its activities, later

moving toward its current state of establishing independent legal status as a research

infrastructure for the arts and humanities. The DARIAH plan for creating an infrastructure of

skills and knowledge, rather than instrumentation or collections, maintains its focus on the

community and its needs, maximizing its ability to get 'below the level of work,' and standing

in opposition to the more common model of a cyberinfrastructure as 'products' -- such as

portals, data centres, archives, digitial archives, digital archive portals, etc. -- rather than what

the research may actually require. To meet this goal, DARIAH takes a somewhat unorthodox

approach to both its facilities and its funding, focusing on 'Virtual Competency Centres' and

national in-kind contributions from digital projects operating at the coal-face of digital discovery.   This focus resists not only the expected norms of what an infrastructure should be, by maintaining its focus not what scholars do being made 'big and fast,' but what the research communities need made accessible, enabling scholarship to begin at the furthest appropriate, established state of knowledge. It also stands in some opposition to the preference among research funding agencies to view infrastructure investment as coterminous with capital investment, eliding the fact that human competence and ideas, rather than buildings and specialist mechanical tools, is often at the heart of what research needs to sustain it. These aspects of the DARIAH umbrella distinguish the service it provides to its community, placing the strongest emphasis in the support it elsewhere in the research process than many of its equivalents.

The CENDARI project was conceived to capitalize upon precisely these opportunities. Its founding vision was based upon the recognition that new methods of technologically enhanced investigation in the humanities supplement, rather than supplant, the tried and true methods of close reading, contextualising primary sources within relevant secondary discourse, and contributing to communities of practice consisting of one's peers.   It recognises the value of the old methods while seeking also to enhancing them. And this is not its only 'balancing act:' CENDARI also stands at the centre of several stakeholder groups: libraries and archives (the old infrastructures), information and communications technologies (the new infrastructures) and end users -- in this case, historians.

The CENDARI project, therefore, was not simply born of an obvious incremental step in the development of research infrastructures for the historical researcher. Instead, CENDARI was required, from the very beginning, to view itself as building on precedent while simultaneously seeking new answers to some of the age-old questions about what researchers 'wanted' from their digital archival research infrastructure. But its position at the centre of several clear stakeholder groups, each of which had a certain tradition and track record even in the development of digital infrastructures, was to become a source of tension, surely, but also a source of inspiration for CENDARI. The ethos of the project, and of the teams it brings together, is one of striving to understand the culture of our collaborators and their contexts, while simultaneously seeking a greater understanding of how work is done,

and could be done, within our own. As such, each of the stakeholder groups in CENDARI has been tasked not just with a set of technical or research deliverables, but also with a 'grand challenge,' representing how they can contribute not only to the shared goals of the project, but also how they can re-envision their own practices through their contribution to a research infrastructure. These challenges are different for each stakeholder group, but represent together CENDARI's goal of producing something more than a portal, more than a tool and more than a set of digital objects.

The recognition of a need to do something different in the CENDARI infrastructure began with the charge to 'integrate digital archival collections.' As mentioned above, libraries and archives have not made the easiest transition to the digital age, a fact which has a long history behind it:

> A look at the history of the Internet reveals a key factor that initial deployment and ramp-up occurred within a tightly scoped community, academic institutions and (primarily defense-related) research labs. The infrastructure had a long percolation period in this context before its subsequent mass popularization. This is quite different than the DL infrastructure work, which from the beginning was motivated by visions of widespread grassroots dissemination inspired by scenarios such as that articulated by then Vice President Gore in his 'schoolchild in Carthage, Tennessee plugs into the Library of Congress' speeches [11]

An early mapping exercise undertaken by the CENDARI project team revealed a further challenge: that there were a large number of collections relevant for the project's two case study areas, such as those of the Imperial War Museum and the French National Library, that were already readily available electronically in high resolution with excellent accompanying metadata. To combine these high value resources and provide domain specific tools for their enhancement and investigation would create an incredible resource for further investigation, but investigation only on already well-known and well-researched collections. At the margins of what was clearly visible, however, was a great sea of material, which was neither readily available nor well-researched, held in archives that for reasons of their own history could not have made the investment in item-level digitisation and cataloguing that the larger centres had done. It became a frightening realisation among the CENDARI partners that the project could incur a credible risk of having a negative influence on the course of historical enquiry, encouraging narrow approaches on well-trodden collections, if it merely integrated the collections already so visible. The concept that arose out of this discussion initiated the

project's focus on so-called 'hidden' archives. Archives could be hidden because of a minimal digital footprint, because of limited staff, because of language issues, and indeed because of contested histories, and it become the first guiding challenge for the CENDARI project to find ways of representing archives without substantial digital collections, perhaps even without adequate digital finding aids, and make their holdings somehow visible enough to the scholar-users of the CENDARI platform that they would be able to understand the potential of a given archive for their research, even if that research itself was still going to move more quickly than for other sites to a need for travel and direct physical work in and with the archives. CENDARI will have passed a major test of its vision to change the paradigm for historical research not when the first archival thesis to be created entirely via CENDARI in a scholar's study appears, but instead when the first scholar decides to invest in learning a new language so as to be able to take advantage of a rich and promising lead discovered through CENDARI.

This ethos of challenge continued with the historians and their input to CENDARI. As the end users of any infrastructure CENDARI might create, it was critical from a social, rather than technical, point of view that historians be centrally involved in the development of the final outputs of the technical development. They also had to be prepared and enthusiastic for the changes in their customary approaches, which the new infrastructure would inevitably open up. The idea to include a case study on the First World War was a very early feature of the project, and advocated by very senior researchers in the field. But, as historian John Horne explains in his volume about that conflict, historians of the era quickly encounter the 'Polish paradox.'[12] Poland was literally and figuratively central in the First World War, and yet we know remarkably little about it and its role. The driving factors behind this are the state of the archives, which were scattered, divided and destroyed (much as the country itself, its borders and its citizens were). The fate of the archives has driven the fate of history, as scholars with limited resources and limited time have focused on the collections available to them in a comprehensible manner, underpinning the tendency to look deeply into national or regional cases (and more deeply into some than others). While the CENDARI project cannot plan to become equivalent to an archive of all material, of all sorts, for all countries with an involvement in a conflict as wide-ranging as the First World War, it does set its sights, as

described above, on providing visibility, access and a wider horizon across the institutions and traditions that hold the archival material. As such, the CENDARI infrastructure will provide historians with powerful tools to overcome the research silos which the research record has encouraged them, by necessity, to develop. But will they be ready? Truly comparative approaches to history will require certain methodological habits to be revisited, not so much in the context of digital versus analogue sources, but far more in the context of how one is able to define the like which one compares with like across institutional, terminological, temporal, spatial and other boundaries. Precision is a hallmark of good scholarship, but what will precision come to mean when new resources with such wildly different provenances are able to be brought together and considered side-by-side? The CENDARI historical team begin their investigations precisely here, where the traditional methods, determined by certain material factors in the records, need to be revised and rethought for the new knowledge ecosystem CENDARI will create.

Defining a grand challenge for the e-Scientists in the CENDARI project was in some ways the easiest to envision but ironically also the hardest. As other observers of cyberinfrastructure have expressed it: 'A tech-centered approach to the challenge of data sharing inclines us toward failure from the beginning, because it leaves untouched underlying questions of incentives, organization, and culture that have in fact always structured the nature and viability of distributed scientific work. Questions of trust loom large here, and run both ways.'[13] The e-Scientists on the team, in spite of their superior expertise in supporting research projects with both technical and humanistic dimensions, could not be allowed to feel too comfortable with that expertise. And this act of discomfiting our colleagues began with language.   Like the term 'portal,' the term 'virtual research environment' has a history of almost a decade of use, and many different kinds of endeavors have produced VREs of widely variant ambition and utility. The project team neither wanted to send the message externally nor encourage any sort of complacency within its own team through the use of a term which had been so broadly defined in practice over the years, in spite of the fact that the key final output envisioned for the project was indeed a virtual space where researchers could find what they needed to undertake their work. From a semantic point of view, a shift was required, and the term which emerged was the 'enquiry environment': it no longer mattered

specifically that it was virtual, only that it would be a place to ask and answer pressing and complex questions. But semantics alone would not safeguard CENDARI from the difficulties tool developers and other computer scientists working with historical domain experts have experienced in other projects. Humanists are commonly referred to as Luddites, but indeed, they neither fear nor resist technology. Word processing, email, Google searches, and specialized tools such as the MLA International Bibliography and Early English Books Online (EEBO) or Eighteenth Century Collections Online (ECCO) are all as much a part of the humanist researcher's toolkit today as they are of the engineer's. What humanists resist, however, is the application of something not fit-for-purpose, or the tweaking phase in which a tool needs to move from fifty per cent effective toward a more reasonable ninety per cent plus. A merciless intolerance for the process of tool making, and for the inevitable partial failures that occur on the road to any success, more accurately describes the traditional humanistic stance to technology, which is necessary only when no better tool exists already in the analogue toolbox. To be successful, the design of the CENDARI environment has to be robust, invisible, and driven by user needs at every level. For this reason, great investment is being made into CENDARI's underlying architecture, its information flows and standards, its data model and ontologies, its integration with existing data management standards, and so on, all the way up through the participatory development of new tools for use within the environment (which, in the current design phase, have ranged from the usual suspects, such as keyword searches, to the quite unexpected, such as TripAdvisor-style local travel assistance for archival visits!).

The final 'grand challenge' for CENDARI exists at the management level, and consists of being thoroughly integrated with the environment around it. Both the 2006 ACLS *Cultural Commonwealth* report[14] and the more recent European Science Foundation report *Research Infrastructures in the Digital Humanities*[15] stress the importance the research cyberinfrastructures for humanities must involve the creation of ecosystems, collaborations and partnerships – between sectors, institutions and individual actors. In this spirit, CENDARI relates not only to the DARIAH infrastructure and other DARIAH associated projects, but also, explicitly or implicitly, to a well-populated environment of digital library and archive projects (such as Europeana[16] and the Archives Portal Europe[17]), to networks of historians in the case

study areas (such as the Society for First World War Studies[18] and the International Medieval Congress[19]), to the physical archives we aim to represent, to the developers of tools and standards we deploy, to cognate projects across the world, such as MESA[20] in the United States, and so on. The project also further progresses its ethos of interconnection through some of its distinctly social, rather than technical development aspects, including summer schools and shorter training events and a programme of transnational access fellowships for scholars, supporting them to bring their projects and perspectives to one of the partner institutions, enriching both their own work and that of the CENDARI team.

With this vision driving its activities, CENDARI seeks to meet its goal of providing scholars with a uniquely rich and flexible virtual space in which to conduct their research, where the breadth of resources, the ease of use and value of the tools available easily overcome any hesitations users may have and compensate for any frustrations they may experience with the limitations of the system. In contrast to other digitisation projects where the vision of the end user and their goals is somewhat broad, CENDARI profits from its focus as a research infrastructure: while other users may indeed find CENDARI provides interesting opportunities to explore historical documents, these are not the core users provided for. Instead, it is the historian for whom primary provision is made, and initial conjectures are already emerging regarding the types of research questions for which CENDARI will provide unparalleled access to insight. Certainly on a generic level, the focus on transnational and comparative history will be manifested through core ontologies and multilingual support to enable the comparison of equivalent documents across countries and institutions. But CENDARI will be able to do far more. The tools layer will allow these connections not only to be suspected, but investigated for potentially unforeseen patterns via the power of geographical, linguistic, temporal and other forms of data visualisation. Data visualisation does not take away the work of the historian, but it can reveal patterns that are otherwise too dispersed or complex to be seen, for example in cases or large corpora of documents which must be read in a linear fashion by the individual scholar, but which can also be 'read' in a distant fashion by a computer query on their contents. In order to promote and inspire such investigations, the CENDARI environment will also offer template research guides, some of which will be historical and thematic (eg 'submarine warfare') while others will be more

historiographical, using proxy data to indicate, for example, the contents of collections known to have been lost through fire, theft or neglect.

It is still too early in CENDARI's development to speak much of its actual achievements, although substantial new understanding of collaborative practice between historical, technical and collections-level standards and practices have already been usefully tackled with a novel input workflow, and some real differences within historical practice itself are being exposed, theorised and incorporated into development practice. But it is still its integration – with other research projects, digital collections and infrastructures – which most distinguishes it, even at this early phase. The research project in the digital humanities struggles with its own mortality: in absence of sustainability dowries or clear pathways for the deposit of functional tools, interfaces or (in many cases) objects, the accessibility of digital project outcomes after their funded period of development ends is often unclear - ironically, this is also the point when their development is, at last, complete and their system as debugged and documented as it is likely to become. This is perhaps regrettable, and perhaps natural for projects, but unacceptable for a research infrastructure. Unlike so many other digital humanities initiatives, however, CENDARI's pathway of assimilation into the wider DARIAH developments has been agreed in principle from its conception, and is firmly in the mind of the project team as the project development proceeds. As such, CENDARI's outputs not only have a reasonable guarantee to providing maximal long term benefit for users, and value for the money invested in its creation, but also a built-in capacity for flexibility and change over the long term. CENDARI may never experience have the 'luxury' of becoming a semi-useful silo, as its components will be automatically vetted for their contribution, on the spectrum from the essential to the interesting, and migrated into a DARIAH format accordingly. In the end, this capacity to sustain its contribution, through future-focused development and carefully constructed partnerships, may be its greatest achievement.

---

[1] European Science Foundation, *Research Infrastructures in the Digital Humanities* (Strasbourg, 2011), 3.

[2] P. Edwards, S. Jackson, G Bowker and C Knobel, *Understanding Infrastructure: Dynamics, Tensions and Design,* http://hdl.handle.net/2027.42/49353, last accessed 16 November 2012, 3.

[3] G. Rockwell, *As Transparent as Infrastructure: On the research of cyberinfrastructure in the humanities,* http://cnx.org/content/m34315/1.2/, last accessed 16 November 2012.

[4] D. Atkins, et. al., *Revolutionizing Science and Engineering Through Cyberinfrastructure: Report of the National Science Foundation Blue-Ribbon Advisory Panel on Cyberinfrastructure,* http://www.nsf.gov/od/oci/reports/atkins.pdf, last accessed 25 January 2013.

[5] European Strategy Forum on Research Infrastructures, *European Roadmap for Research Infrastructures, Report 2006* (Luxembourg, 2006).

[6] J. Unsworth, et. al., *Our Cultural Commonwealth. Report of the American Council of Learned Societies Commission on Cyberinfrastructure for the Humanities and Social Sciences, 2006,* http://www.acls.org/cyberinfrastructure/cyber.htm, last accessed 12 November 2012.

[7] European Strategy Forum on Research Infrastructures, *European Roadmap for Research Infrastructures, Report 2006*, 5

[8] http://www.nines.org

[9] http://www.projectbamboo.org

[10] Edwards, Jackson, Bowker and Knobel, *Understanding Infrastructure*, 17.

[11] C. J. Lagoze, *Lost Identity: The Assimilation of Digital Libraries into the Web*, Ph.D. (Ph.D. thesis, Cornell University, 2010). Cited in G. Rockwell, *As Transparent as Infrastructure: On the research of cyberinfrastructure in the humanities,* http://cnx.org/content/m34315/1.2/, last accessed 16 November 2012.

[12] J. Horne, *Commemorating the Centenary if the Great War and the Division of Ireland: A European Perspective*, http://www.ucd.ie/ibis/filestore/wp2011/106_horne.pdf, last accessed 16 November 2012.

[13] Edwards, Jackson, Bowker and Knobel, *Understanding Infrastructure*, 32.

[14] Unsworth, et. al., *Our Cultural Commonwealth*.

[15] European Science Foundation, *Research Infrastructures in the Digital Humanities*.

[16] www.europeana.eu/

[17] www.apex-project.eu/

[18] http://www.firstworldwarstudies.org/

[19] http://www.leeds.ac.uk/ims/imc/

[20] http://www.dlib.indiana.edu/projects/mesa/