Sustainability of Digital Humanities Projects as a Publication and Documentation Challenge

<table>
<thead>
<tr>
<th>Journal:</th>
<th>Journal of Documentation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID</td>
<td>JD-12-2019-0232</td>
</tr>
<tr>
<td>Manuscript Type:</td>
<td>Article</td>
</tr>
<tr>
<td>Keywords:</td>
<td>digital humanities, project documentation, project management, sustainable publishing models, humanities research, collaborative research</td>
</tr>
</tbody>
</table>
Sustainability of Digital Humanities Projects as a Publication and Documentation Challenge

Structured Abstract

Purpose – This paper proposes a new perspective on the enormous and unresolved challenge to existing practices of publication and documentation posed by the outputs of digital research projects in the humanities, where much good work is being lost due to resource or technical challenges.

Design/methodology/approach – The paper documents and analyses both the existing literature on promoting sustainability for the outputs of digital humanities projects and the innovative approach of a single large-scale project.

Findings – The findings of the research presented show that sustainability planning for large-scale research projects needs to consider data and technology, but also community, communications and process knowledge simultaneously. In addition, it should focus not only on a project as a collection of tangible and intangible assets, but also on the potential user base for these assets, and what these users consider valuable about them.

Research limitations/implications – The conclusions of the paper have been formulated in the context of one specific project. As such, it may amplify the specificities of this project in its results.

Practical implications – An approach to project sustainability following the recommendations outlined in this paper would include a number of uncommon features, such as a longer development horizon, wider perspective on project results, and an audit of tacit and explicit knowledge.
Social Implications – These results can ultimately preserve public investment in projects.

Originality/value – This paper supplements more reductive models for project sustainability with a more holistic approach that others may learn from in mapping and sustaining user value for their projects for the medium to long terms.

Keywords: digital humanities, project documentation, project management, sustainable publishing models

Article Type: Research paper
Introduction

In spite of the increasing prominence of the term ‘sustainability’ in both academic discourse and in the research funding and policy circles that engage with the digital humanities, there is little consensus as to what exactly the word means, or what exactly sustainability should achieve. According to the Oxford English Dictionary, the term can refer to two parallel, but different capacities: to be maintained ‘at a certain rate or level,’ (as in sustainable growth) and ‘to be upheld or defended as valid, correct, or true’ (as in a sustainable argument).

The first of these senses, while closer to what one might perhaps think of as the challenge faced by digital humanities projects, does not quite capture the imperatives that are being applied to such projects’ research results. They are, in a general sense, expected to be available for the long term and accessible to potential users (hence the rate and level of maintenance) but there is also an implication that to be sustainable can also mean that this should be possible without further requirement of investment or resources from the original funder of the work. In addition, such a vision of sustainability maintains a narrow vision of scholarship as able to be adequately captured through certain kinds of products, such as the monograph or article, but perhaps also the website or dataset. This perspective is useful to highlight the publication challenges that are raised by these still new forms of digitally-enhanced research and communication, but it misses emerging recognition that much of the value of scholarship must be captured in its processes, not merely in its products. Digital humanities approaches have not merely changed the way in which scholars express their findings, but also the manner in which they work, the tools they build, the disciplines they connect to. In particular in multi-year, collaborator projects, this layer is liable to go undocumented, and therefore unrecognised, in spite of the ground-breaking methodological or other process-based work that may have occurred.
This tacit understanding of sustainability (in the context of the digital humanities) is based on the mental model of the book: once complete, it is relatively stable and (within the normal expectations for environmental conditions) likely to be available and discoverable through a library for the long term. Sustaining the knowledge generated by research projects does not, unfortunately, abide by these simplistic rules, however, and remains an ongoing challenge within the wider digital humanities ecosystem. Even the recognisable outputs of such projects can tend to become ‘orphaned,’ unmaintained, and eventually disappear in a context where ‘technological optimism was somewhat higher and security risks somewhat lower than they are today’ (Smithies et. al., 2019), leading some to claim that ‘…given the current state of digital preservation methods, sustainability is virtually impossible. Digital material is simply not configured for preservation beyond a few years, perhaps a few decades at most’ (Green et. al., 2017, p.15). This is particularly the case for projects that can be billed as research infrastructure, where the scale, complexity and indeed the overarching aim of the project to serve a potentially still emerging community makes their continued accessibility even more important, and even more difficult. The experience of the Bamboo Project is a good example of how attending too late to the challenges of sustainability after the end of a funding period can lead to the ‘failure’ of an ambitious project (Dombrowski, 2014).

Expectations based on centuries of experience with libraries and books and scholarly arguments that were largely univocal and largely argument-based, have formed the ways that we speak about sustainability as well as the ways in which we try to create it. These models, and others that have attempted to displace them, often lead projects down ineffectual paths, laden as they are with expectations and assumptions that cannot necessarily be translated effectively into a digital ecosystem. This problem intersects with some of the key concerns of contemporary scholarly communications in the humanities: new forms of scholarly process and output are proliferating faster than systems can be built to recognise and preserve their contributions, and to meet for them the traditional aspects of scholarly publishing, such as dissemination, registration,
archiving, and accreditation. As a result, project level interpretations of value and durability are driving decisions in a number of directions, each of which may based on a large number of tacit judgements regarding what scholarship is, and what it does.

Implicit Conceptions of the Project: Implicit Definitions of Sustainability

As Digital Libraries have the longest tradition in the preservation of digital objects and their related metadata, the route they have taken in relation to the topic of sustainability has been formative for later work. Already in 2003 the Council on Library and Information Resources published a booklet entitled ‘A survey of digital cultural initiatives and their sustainability concerns’ (Zorich, 2003), in which the author, Diane M. Zorich, introduced a number of factors that need to be considered when discussing sustainability, elements that she characterised as ‘threats’ to the development of digital projects, such as continuity of funding, data preservation (including choice of standards and longevity of data) and the need for flexibility in order to change business models, when necessary. More recently, libraries also have expressed a keen interest in incorporating digital humanities centres and research activities as a way of ensuring mutually beneficial development of methodological support, rather than promoting a potentially problematic or unbalanced model service provision (see Montoya, 2016; Posner, 2013; Siemens et. al. 2011). Needless to say, collaboration to support the sustainability of project outputs features centrally in these discussions (see Irvine, 2015; Green et. al., 2017; Vinopal and McCormick, 2013; Sweeney et. al., 2017; Montoya, 2016; and Schaffner et.a.; 2014). A comprehensive review of literature published in LIS journals about the topic of sustainability (Eschenfelder et. al., 2019) shows that this interest remains strong, but also belies a number of the limitations on the understanding of a digital humanities project and its contribution, discussed above. Sustainability was generally understood as ‘the continued operation of a collection, service, or organization
related to digital libraries, archives, and repositories over time’ (Eschenfelder et. al., 2019), but was also generally conceptualised as simply as ‘a value (similar to inclusiveness or efficiency) that ought to motivate action.’ (Eschenfelder et. al., 2019).

This literature and others reviewed in the course of developing the CENDARI project sustainability approach led to the conclusion that there were in fact a large number of definitions of the term ‘sustainability’ in circulation, most of which envisioned the project or its assets according to a different model or metaphor. While each of these models (which will be discussed in the following sections) represents a frame of reference that contributes greatly to the overall understanding of digital projects and their contribution, the manner in which they seem to operate as silos may be a hindrance as well, losing some of the richness and context that makes the digital project such a strong contributor to contemporary humanities research.

MODEL 1: SUSTAINABILITY AS PRESERVATION OF DATA AND METADATA

It is notable in this respect that when it comes to discussing sustainability, research and reflection coming out of the digital library community converges on a few key themes. One of the most prominent of these perspectives views sustainability as a technical function, related to the need to preserve digital objects, data, metadata and tools. For example, an early document created within the University of North Texas Library (Alemneh, Hastings and Hartman, 2002), indicates a series of steps to be taken for a ‘healthy’ data sustainability plan, including a Life Cycle Assessment of the Digital Resources, Draft of a Metadata Architecture, Metadata Creation Workflow, and Metadata Creation Tools. All of these elements are hugely useful contributors to the sustainability of the work of the digital library, but are also indicative of a relatively narrow conceptualisation of the impact of the digital library, viewing the data and collections of the library as the sole object of a sustainability process. Other work reflects this same conflation in different ways. The IFLA Manifesto for Digital Libraries (IFLA, 2010) for example, also focusses on data as the key element to be sustained, but through the lens
of interoperability, rather than preservation. This theme continues to this day, with one library-based author speaking of sustainability in the context of his work as applying to ‘… a sub-genre of digital scholarship that primarily depends on the creation of relatively static, combinatory content.’ (Montoya, 2016, p. 606)

These examples of sustainability-related discussions reveal a relatively traditional viewpoint, in which a digital project is viewed not so much as a new kind of scholarly initiative, but an institution equivalent to its bricks and mortar precursor, built to house and maintain collections, which does not need to prove and define its value on its own term, but to manage threats to its fledgling existence.

MODEL 2: MODEL 3: SUSTAINABILITY AS MAINTENANCE OF AN ORGANISATION OR INSTITUTION

Another conceptualisation of the digital project or infrastructure development that arises in the literature around sustainability is of the project as a form of organisation or institution. Often such approaches are focussed very strongly not so much on data as the key asset to guarantee access to at a certain rate or level, but on creating a ‘business model’ for maintaining the user-facing services and the organisation (including staff) created (see, for example, Maron, Kirby Smith and Loy, 2014). The Archives Portal Europe project, for example, decided to establish itself after a funded period of development as a Foundation, as this organisational continuity was considered to be critical in order to maintain the vital functions of the project and to further connect with other similar initiatives. This creation of a Foundation was foreseen in APEnet project sustainability report (Andrén et al., 2012), which described it as the best instrument by which to mitigate the risks and ensure a future of sustainable growth: this mechanism was viewed as a way to facilitate activities such as international cooperation and collaboration, sustainability of the digital archival infrastructure and the assessment of use cases of digital services.

This second mode of conceptualisation is also evident in a the research undertaken by JISC and Ithaka S+R in 2008, which was published in the report
‘Sustaining Digital Resources: an on-the-ground-view of Projects Today’ (Maron and et al., 2009), and also later into a ‘Sustainability Implementation Toolkit’ (Maron and Pickle, 2014). The analysis of the JISC/Ithaca report depicts a scenario where project management strategies are crucial in the sustainability of a project after the end of its funded period: dynamic and engaging leadership, identification of a value proposition, creative cost management, and cultivation of diverse sources of revenue are all factors that reflect the well-being of a digital project under this model.

MODEL 4: SUSTAINABILITY AS ENSURING TECHNICAL ROBUSTNESS FOR A TOOL OR TECHNICAL PLATFORM

A further, more recently emerging model views the project as a tool or technical platform, with the resulting sustainability proposal primarily taking into consideration issues and practices such as migration and hosting of elements such as the repositories where the data are stored and continued maintenance of work environments and specific tools. The TextGrid Project, for example, has been hugely successful in ensuring that its platform has remained accessible over a long period, continuing to make its services available to users. At its best, this approach results in a broad focus on software durability, documentation of processes and the modularity of services (Buddenbohn et al., 2015).

Similar to TextGrid, the DM2E (Digital Manuscripts to Europeana) project invested significant effort in the development of a model to sustain its activities and tools after project close. The goal of the project was to build tools and the communities to enable humanities researchers to work with Manuscripts in the Linked Open Web. The project worked on different levels of knowledge creation and transfer in order to enable scholars - as well as third party projects such Europeana - to reuse the content they gathered and to enrich it with annotation tools such as Pundit. In 2014 the project released a deliverable on the sustainability of the project, entitled ‘Result Transfer Plan’ (Goldfarb et al., 2015). This plan focussed largely on transferring technical management for the data, tools and services developed by the project to the Europeana Digital Library to manage.
MODEL 5: SUSTAINABILITY AS A USER ISSUE

All of these perspectives originate from that of the resource creator, however, and it is not surprising that sustainability statements formulated by people who build things would tend to focus on ensuring the availability of the things they have designed, built and can (to a certain extent) control. But sustainability has a user side as well, and it is in the nature of short-term, project-based funding that project trust, visibility and tuning to the needs of these users, out there, in the wild, may still be in its infancy as the project team adjourns. The importance of these intangible assets cannot be underestimated, however, as these are the elements that will assist project results to be used and more importantly reused. The LAIRAH project provides an excellent example of how to access this perspective from the outside, describing what factors use analysis reveal as key enablers for digital projects and tools to be found and adopted by users (Warwick et al., 2008). The LAIRAH findings seem out of step with the high-level statements projects tend to make in their sustainability plans, however, and suggest another model for the sustainable project, in which the communication and branding of the project is a key element of its success, and handover to the user community is a part of the thinking behind the long term use model.

While each of these models has strengths and weaknesses, what is often missing from them is the conceptualisation of where the vision of the project creator intersects with the needs of the future user of the infrastructure, in particular when that user is likely to have a different set of skills, needs and goals from those of the original team. This is what we would call a re-use, or use value, paradigm, taking in both publication and documentation aspects, but it takes a certain amount of imagination and perhaps even courage to realise, as will be described below.

The Project as One and Many
The reliance upon metaphors able to capture the essence of a digital project and align it with a known field of objects and activities appears as a common thread in the writing and thinking about sustainability. This alignment, once it is achieved, makes the digital development less liminal, able to be imagined into the future as maintained at a certain rate or level, and able to imply, by the nature of its alliance with collections, organisations, software, brands or other model, a definition for what sustainability should mean.

These metaphors, however, become boundary objects in the sustainability discourse, in two senses. First of all, they operate in the sense that Star and Griesmer (1989) have described, as 'objects which are both plastic enough to adapt to local needs and constraints of the several parties employing them, yet robust enough to maintain a common identity across sites...weakly structured in common use, and become strongly structured in individual-site use [with] different meanings in different social worlds but their structure is common enough to more than one world to make them recognizable, a means of translation' (393). By implying an identity that may or may not encompass their true scale and scope, the metaphors that tend to underpin sustainability planning enable understanding across intersecting social worlds (such as between researchers and funders), but they may also betray the very developments they seek to make comprehensible, or, indeed, to transcend the boundaries of. This is the second manner in which the idea of the boundary object can be applied: literally, these conceptualisations place boundaries upon the development, which may or may not encompass the full richness of the projects and the elements to be taken into consideration in the sustainability planning. The metaphors are, in fact, synechdoches, reflecting the project team's desire to ensure their results are sustainable, but at the cost of reducing some of the richness of the original development endeavor.

A successful approach to sustainability for projects, and in particular for projects claiming to develop research infrastructures, needs to be more comprehensive, therefore, not just considering data or technology, community, communications or processes, but in fact all of them simultaneously. In addition, it should focus on a project as a collection of
tangible and intangible assets, but also from the perspective of the potential user base for these assets, and what these users consider valuable about them. Finally, it needs to take its definition of sustainability from somewhere other than the norms of an imposed model adapted via a metaphorical transfer of values and characteristics. Sustainability must be underpinned not just by an understanding of ‘what’ and plan for ‘how’, but by solid conceptions of ‘why’ and ‘for whom.’

The CENDARI Project Approach to Sustainability

Faced with this landscape of good, but limited, practice, the Collaborative European Digital Archival Research Infrastructure project (known also as CENDARI) decided to implement a different approach, one based upon a definition of sustainability as ‘optimised for reuse’ and of the project as not just one thing, but as many simultaneously, all of which may have reuse value for a particular community or in a particular situation. CENDARI was an excellent test case for this approach, having, as it did, both a significant scale and a number of key assets that would allow it to test its planning. CENDARI had in fact committed to the production of a Business Plan in its original description of work, but as the project team began their work, it quickly became apparent that CENDARI was not so much a ‘business’ (an organisation to support commercial transactions between suppliers and customers), but rather many other things – a set of communities, a set of tools, a knowledge base, a set of resources – none of which met the definition of a business.

Another contributor to the early need to reconsider the proposed approach to sustainability within CENDARI was the assumption that the project’s relationship with a durable European research infrastructure consortium, the DARIAH ERIC, would guarantee that its results would be maintained. But a complex technical infrastructure cannot simply be frozen in time and expected to continue to meet evolving needs. It emerged early in the project that the communities that built CENDARI would need a more
active platform to ensure this, and the assets produced would need a broader base than
even DARIAH could provide in order to maximally benefit its potential users. It is from
these two realisations that the need to question and indeed redefine sustainability was
recognised within the project, and adopted as a piece of meta-research running in parallel
to development of the project's platform, data and research cases. Although both the
generic process and the specific actions resulting from this research match on some level
choices made for the CENDARI development, they also reflect the reality, identified by
Joris Van Zundert, of the ‘fluidity’ of research infrastructure, caught up in both the digital
information lifecycle and the creation of knowledge by end users, as well as the software
components (Van Zundert, 2012, 179).

The CENDARI sustainability plan therefore looked beyond the previous work
described above in a number of ways: first, it presented a view of sustainability as a
process, rather than a state, which begins with project conceptualisation and ends far
after project close; second, it views the end goal of sustainability not as stasis, but as
transformation, and reuse; finally, it views the project not as an organisation or technical
platform, but as a collection of tangible and intangible assets with potential value to other
users. As such, the plan proposed a comprehensive understanding of infrastructural
assets, and a model for sustainability that considered not only the most obvious end
product of the project (in CENDARI's case, a virtual research environment), nor even just
the project’s tangible, reusable and customisable components - such as data and tools -
but also the project’s intangible achievements, such as the know-how and the expertise
developed during its duration, and the community of users that had been built through the
project’s testing and outreach phases.

The CENDARI Assets and Multilevel Sustainability Action Plan
The first component of the sustainability planning process, an audit of assets and reuse scenarios, served to resist the common impetus to base sustainability measures on a conceptualisation of the project that might not be comprehensive, agreed or indeed explicitly recognised. This audit (which included both internal mapping and a stakeholder validation meeting) refined project-level understanding of the full range of assets the project had generated and how they could be maintained, shared and indeed passed on to its key users for further development. This process identified seven categories of assets as most likely to find future usage, each of which posed unique challenges in how they could be captured, made visible and sustained. It was one of the greatest challenges of the CENDARI sustainability planning process to ensure that for each of these categories a solution to the challenges of sustainability was implemented, making them findable and reusable in a contextualised manner, and preserving them in multiple formats and multiple locations as needed. For each asset type, the audience for potential future use would be different, and therefore the solutions proposed generally were as well.

A brief overview of these seven categories follows. Although other projects would almost certainly have a different mix of assets in their portfolio and place differing emphasis on their importance, this list is indicative for the kinds of categories a project should expect, and some of the measures that can sustain them.

The CENDARI portal was the most visible of its assets, representing the final synthesis of the project’s activities and its main point of access. For many projects, this would be where sustainability planning would not only begin, but end. CENDARI approached this sustainability challenge via a three pronged strategy, guaranteeing 3 years of access through the German arm of the DARIAH infrastructure but also ensuring new communities and new approaches would be recruited to continue development. To this end, the portal was also prepared and preserved as a virtual machine, known as ‘CENDARI in a box,’ which is available for research communities that would like to create a new installation of the project environment with their own data or tools. In this way,
reuse of the technology underpinning the portal was separated from the portal itself, in recognition that although its customised and dedicated ‘front door’ might seem the most important element produced by any project, it is also generally the hardest to maintain, and the most technically fragile over the medium to longer term.

The portal was not only useful in its complete final form, however, but also as a collection of unique services, tools and components optimised to support digital humanities research. This possible tactical reuse of the project outcomes was foreseen from the beginning, and a very modular, service-oriented architecture was adopted for the project. The tools therefore required a sustainable pathway outside of the portal as well, although connecting tools with potential user bases is a constant challenge. The software community practice of sharing software through GitHub, which Spiro and Smith estimate to be used by around 20% of digital humanists (Spiro and Smith, 2016) was adopted, but further awareness raising through dissemination and communication was also required to ensure the maximal future use for the tools.

CENDARI holds a lot of data from different sources and in different formats, some unique to the project, others well signposted elsewhere, all with different requirements and expectations for sustainability. This has been its legacy as a project seeking to reuse archival data for historical research, where the culture and ability to share data was unevenly developed within the partner archives. The CENDARI data portal gives access to this data, and the project’s data agreement and license were developed with DARIAH as a co-signatory, ensuring that a permanent contact point for use and reuse questions would be maintained. But DARIAH is not well-known as a data provider or source, and this solution alone may not maximise visibility and reuse. Therefore a redeposit protocol for unique data with an external trusted source was also facilitated. Institutions were also offered the opportunity to reingest their own, transformed data, but in most cases the institutions did not have an appropriate workflow in their side to make use of this, which makes it all the ore important that projects consider their data production as separate from both the technology that might deliver it and the publications that might feature it. A
separate deposit of these assets as research, rather than cultural, data is therefore a more sustainable approach to take under current conditions, where numbers of repositories and finding aids for research data are growing and consolidating.

The CENDARI Archival Research Guides, which combined the elements of an archival finding aid with a scholarly interpretation in a dynamic, multimedia format, existed as a particular subset of the data held unique to CENDARI, but their status as both primary and secondary research sources justified their consideration as an asset class in themselves, in particular because of the manner in which they challenged existing norms of publication, communication and evaluation in the discipline of history. As extended and enhanced publications, incorporating analysis, links to data sources, multimedia objects, and links to project ontologies, these guides needed to be delivered within the project portal to be used to their full potential. But to sustain these unique works of scholarship only in that format would again potentially limit their visibility. They are therefore offered in several export formats, as well as becoming the focus of both a review publication and a research paper in a mainstream (not digital) historical journal. In this way, their contributions to scholarship can be recognised as independent from the format in which they have been delivered, while the affordances of that format can still be harnessed.

Given how particular many of the project experiences in building for the DH community had been, a specific audit of CENDARI’s tacit knowledge was also undertaken, and several white papers and process oriented toolkits emerged from the project on the foot of this (including, for example, the ‘White Book of Archives,’ (Beneš et al., 2015) documenting the project’s experience of federating highly heterogenous data from traditional collection holding institutions).

Perhaps the least easily defined and sustained aspects of the CENDARI project were the communities - mixed and homogenous groups of historians and other humanistic scholars, collections experts and technologists - it brought together and fostered for four years. Interconnectivity between cognate projects was a key resource for facilitating this, as some communities will have interests across these projects, and
networks can and should be shared. Some voices in infrastructure studies do recognise the importance of considering this layer, encouraging projects to view ‘technical infrastructure as a human and community asset in need of maintenance and support, rather than a technical artefact in need of service management’ (Smithies et al., 2019). But adopting this perspective that puts the community first remains a challenge, which many projects face with the development of more technical infrastructure, such as the Modernist Commons focus on sustaining its community via ‘Web-accessible repository of digitized resources, a Web-based workbench of digital editing tools, and a customizable publication interface’ (Irvine, 2015, p. 5). The CENDARI team felt that such an approach would not be appropriate, however, given the place of the digital in the workflows of the community it served, and the nature of the interactions and consensus required to ensure true sustainability. For this reason, the project decided to use another DARIAH mechanism, the Working Group, to provide a structure for continued development of project concerns and assets, monitoring the value it is providing at regular time steps and taking action to maintain or indeed increase it.

**High-Level Principles to Foster Sustainability**

In analysing the full possible range of reuse scenarios for these classes of assets, the CENDARI project developed a set of five high-level principles to secure the reuse potential of any digital project. These are:

1. **EMBED INTO PARTNERSHIPS and NETWORKS:** Users will need to know how any new tool or service fits in with the existing landscape they know. Working with other projects and organisations will help you to build your networks as well as your credibility, as well as giving you possible partners at the end of the project to take on and further development of your data sets or tools. By sharing widely and embedding in networks, you can build a solid basis for a core group to continue oversight of the project
results: this is easier to facilitate if a new umbrella identity can be established. The DM2E
transfer of results to Europeana is an excellent example of how this principle can be
delivered upon.

2. DEVELOP WITH MODULARITY AND OPENNESS: Designing a digital project
in a way that allows single elements to be reused and implemented by other research
infrastructures and projects will add a rich further layer to the possible scenarios for their
reuse. In addition, projects should use common platforms, such as GitHub for software or
an open repository for reports, and known and accepted standards in their work, as
aspect also recognised in other projects (Vinopal and McCormick, 2013, p. 31). These
may be drawn from a number of relevant communities (TEI for scholarly editions, EAD or
ISAD for archival material, etc.) but maintaining reference to a larger context for the
organisation of data should be viewed as an essential element of its potential reusability.

3. DOCUMENT, VALIDATE and SHARE: Knowledge sharing across the project
and beyond should be not just the norm, but a process projects actively consider and
invest in. This should be implemented not just at the level of project heads, but across
the cognate skills areas within and between projects, with communications people
building relationships with other communications people, for example. Mapping project
structures for this purpose against each other should occur as a matter of course. Project
publications should cover both scientific and management aspects of the project, with
different audiences in mind for different pieces, and, where project publications challenge
community norms, a source of external validation should be identified for them via open or
tailored peer review, or consultation mechanisms. To be successful, this tenet should be
interpreted as broadly as possible: the CENDARI project was in the very fortunate position
to have a preferential relationship with a durable, transnational research infrastructure, but
a committed institution, network, or indeed a commercial publisher may also play a
significant role for a project, as in Laura Mandell’s experience partnering with Gale for the
improvement of the EEBO transcriptions: when asked ‘What is your sustainability plan?’
for the grant deliverables you are producing for the eMOP project, I would answer, “Gale. Gale is my sustainability plan.” (Mandell and Grumbach, 2015, p.3)

In addition, every project should have a thorough data management plan, explaining what data it will collect or create data and where it will be left, including as well any services the project will create to manage this data and how they will be sustained. Technical documentation of the tools and their integration should be openly available. Finally, a ‘Tacit Knowledge Audit’ should be built into the project so that applied knowledge and lived experience does not get lost with project closure, but can rather be shared explicitly in an appropriate format. In particular assets such as usage policies, management plans, cooperation agreements and the like can be of great value to successor projects.

4. MAINTAIN CLARITY OF PURPOSE and SIMPLICITY OF ACCESS: This can be facilitated by designing, for example, a data ingestion cycle that is capable of being rolled out as an easily managed service, providing end users with simple instruments or forms to contact the project’s team, to add content, report bugs and query usage of tools. Maintaining a consistent central communication point, even after active project close, and ensuring that research work can be accessed reliably in a variety of easy to find, relevant formats and locations (such as a trusted digital repository or traditional journal) will support this objective as well. Finally, a project will benefit from being clear about what it has created or collected, what its value is, and for whom.

5. APPROACH SUSTAINABILITY AS A PROCESS, NOT AN END PRODUCT. The CENDARI sustainability planning process was comprised of a series of stages, stretching from pre-planning and mature project to closure and post-project actions, each of which contributed to the overall, holistic sustainability strategy. This cycle was intended to counteract the natural impetus within projects to view sustainability as a concern only for the final phase of the project, rather than one to be integrated into the project’s development and even its conception. Implementing such a process requires focus and resources, but results in a far better and more grounded hand-over phase for the project.
from a funded period to a post-funded period. To capture the benefits of this process, project teams should make sustainability a part of their initial proposals, including at a minimum a sustainability meeting at the project mid-point and a sustainability report as a project outcome.

From a technical standpoint, projects should seek to guarantee the availability of their basic technical formats for 3 to 5 years after close. This may seem short, but it is not necessarily reasonable to expect or support a longer window, as beyond that time it would need to be significantly updated anyway. Committing to a defined, shorter span will encourage early and sustained dialogue between stakeholders. Where projects continue to prove their utility to new and old user communities within and beyond this basic window, successor funding projects to focus on new user groups or new functionalities should naturally come to be sought, ensuring the project maintains the ability to be ‘evolving and involving.’ Where the usage becomes more sporadic and specialised over time, an alternative should be sought, such as clearly and richly documenting the user interface and tools, such as search queries, but ultimately extracting the data layer from the software designed to deliver it for deposit in a long term, stable, open resource, a process Geoffrey Rockwell and his coauthors refer to as ‘Burying a Dead Project’ (Rockwell et al., 2014). The model offered by the Kings Digital Lab to sustain digital projects is groundbreaking in this way, offering as it does a choice of hosting options under a service level agreement, via migration to another institution, or as archived data (Smithies et. al, 2019). Even under such an agreement, however, a project should be sure to think about what might happen in the future, for example when a new host, in particular a commercial one, looks to migrate projects. As Laura Mandell reflected in the context of her work with Gale, ‘We do need to have in our contracts with such vendors that, when the company or its successors no longer wishes to sell and sustain a digital catalogue, it must pass it to libraries for dark archiving.’ (Mandell and Grumbach, 2015, p3)
Conclusions

The experiences of the CENDARI project illustrate how sustaining the results of a digital research project involves so much more than ensuring that a robust technical infrastructure has a long-term hosting solution in place. At the other end of the spectrum, however, it is also counterproductive to view every such project as so different as to require a unique solution. While the circumstances, networks and conditions for sustainability may be unique to each project, the goals, frameworks, best practice experiences and tools underpinning sustainable work will be very much the same across domains and projects. It is therefore one of the great achievements of the CENDARI project to have been able to propose not only a new definition of sustainability, with the definition of being ‘optimised for reuse,’ but also to have been able to discuss and experiment with the mechanisms underlying the principles outlined above.

These principles should be of use not only to project coordinators, but also to project funders and to the wider digital humanities community. Although the movement toward Open Science and FAIR Data has largely gained momentum only after the CENDARI project was conceptualised, and indeed even after its launch, the approach the project took to documenting its many potential contributions and making them explicit very much reflects these emerging scholarly values. By agreeing to the principle that sustainability as reuse to be addressed from the moment of submission (if not conceptualisation) of a proposal, the digital humanities community can continue its move beyond silos and toward richer fields of knowledge exchange and development. Three years after the conclusion of its active phase in 2016, the CENDARI project continues to act as a quiet resource for the inspiration and instigation of individual researchers and projects that follow it. There is of course more to do: recognition of the unique contributions and requirements of the scholarship produced by such projects lags behind our ability to produce it, as does a more widespread recognition of how humanities
research that is collaborative, digitally-enabled, open-ended, and performed ‘at scale’ outside of traditional institutions and networks can be designed, developed, and validated holistically. We hope the experiences of the CENDARI project can make an ongoing contribution to this cultural change within our communities and disciplines.

**Funding**

This work was supported by the European Commission’s Seventh Framework Programme [grant number 284432].
References

First Monday 7(8).

Andrén, P., et. al. (2012), APEnet Sustainability Report
http://pro.europeana.eu/files/Europeana_Professional/Projects/Project_list/APEnet/Deli

2017).

Beneš, J., et. al. (2015), The CENDARI White Book of Archives,


Dombrowski, Q. (2014), ‘What Ever Happened to Project Bamboo?’ Literary and


Goldfarb, D., et. al. (2014), DM2E Results Transfer Plan,
http://dm2e.eu/files/D5.10_1.0_Results_Transfer_Plan_140730.pdf (accessed 24 July
2017).

Green, H. et. al. (2017), Scholarly Commons Digital Humanities Needs Assessment Study
(University of Illinois, 2017)
<https://www.ideals.illinois.edu/bitstream/handle/2142/100081/Final_DHNeedsAssess
ment_Report.pdf?sequence=2&isAllowed=y>

IFLA Governing Board (2010). IFLA/UNESCO Manifesto for Digital Libraries,


