



Royal Netherlands Academy of Arts and Sciences (KNAW) KONINKLIJKE NEDERLANDSE AKADEMIE VAN WETENSCHAPPEN

Sustainability of digital humanities projects as a publication and documentation challenge

Morselli, Francesca ; Edmond, Jennifer

published in

Journal of Documentation
2020

DOI (link to publisher)

[10.1108/JD-12-2019-0232](https://doi.org/10.1108/JD-12-2019-0232)

document version

Early version, also known as pre-print

[Link to publication in KNAW Research Portal](#)

citation for published version (APA)

Morselli, F., & Edmond, J. (2020). Sustainability of digital humanities projects as a publication and documentation challenge. *Journal of Documentation*. <https://doi.org/10.1108/JD-12-2019-0232>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the KNAW public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain.
- You may freely distribute the URL identifying the publication in the KNAW public portal.

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:

pure@knaw.nl



**Sustainability of Digital Humanities Projects as a Publication
and Documentation Challenge**

Journal:	<i>Journal of Documentation</i>
Manuscript ID	JD-12-2019-0232
Manuscript Type:	Article
Keywords:	digital humanities, project documentation, project management, sustainable publishing models, humanities research, collaborative research

SCHOLARONE™
Manuscripts

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.

1
2
3 **Social Implications** – These results can ultimately preserve public investment in projects.
4

5 **Originality/value** – This paper supplements more reductive models for project
6 sustainability with a more holistic approach that others may learn from in mapping and
7 sustaining user value for their projects for the medium to long terms.
8
9
10

11
12
13 **Keywords:** digital humanities, project documentation, project management, sustainable
14 publishing models
15
16

17
18
19 **Article Type:** Research paper
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

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, collaborative 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.

1
2
3 This tacit understanding of sustainability (in the context of the digital humanities) is
4 based on the mental model of the book: once complete, it is relatively stable and (within
5 the normal expectations for environmental conditions) likely to be available and
6 discoverable through a library for the long term. Sustaining the knowledge generated by
7 research projects does not, unfortunately, abide by these simplistic rules, however, and
8 remains an ongoing challenge within the wider digital humanities ecosystem. Even the
9 recognisable outputs of such projects can tend to become 'orphaned,' unmaintained, and
10 eventually disappear in a context where 'technological optimism was somewhat higher
11 and security risks somewhat lower than they are today' (Smithies et. al., 2019), leading
12 some to claim that '...given the current state of digital preservation methods, sustainability
13 is virtually impossible. Digital material is simply not configured for preservation beyond a
14 few years, perhaps a few decades at most' (Green et. al., 2017, p.15). This is particularly
15 the case for projects that can be billed as research infrastructure, where the scale,
16 complexity and indeed the overarching aim of the project to serve a potentially still
17 emerging community makes their continued accessibility even more important, and even
18 more difficult. The experience of the Bamboo Project is a good example of how attending
19 too late to the challenges of sustainability after the end of a funding period can lead to the
20 'failure' of an ambitious project (Dombrowski, 2014).
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40

41 Expectations based on centuries of experience with libraries and books and
42 scholarly arguments that were largely univocal and largely argument-based, have formed
43 the ways that we speak about sustainability as well as the ways in which we try to create
44 it. These models, and others that have attempted to displace them, often lead projects
45 down ineffectual paths, laden as they are with expectations and assumptions that cannot
46 necessarily be translated effectively into a digital ecosystem. This problem intersects
47 with some of the key concerns of contemporary scholarly communications in the
48 humanities: new forms of scholarly process and output are proliferating faster than
49 systems can be built to recognise and preserve their contributions, and to meet for them
50 the traditional aspects of scholarly publishing, such as dissemination, registration,
51
52
53
54
55
56
57
58
59
60

1
2
3 archiving, and accreditation. As a result, project level interpretations of value and
4 durability are driving decisions in a number of directions, each of which may based on a
5 large number of tacit judgements regarding what scholarship is, and what it does.
6
7
8
9
10

11 **Implicit Conceptions of the Project: Implicit Definitions of** 12 **Sustainability** 13 14

15
16
17 As Digital Libraries have the longest tradition in the preservation of digital objects
18 and their related metadata, the route they have taken in relation to the topic of
19 sustainability has been formative for later work. Already in 2003 the *Council on Library*
20 *and Information Resources* published a booklet entitled 'A survey of digital cultural
21 initiatives and their sustainability concerns' (Zorich, 2003), in which the author, Diane M.
22 Zorich, introduced a number of factors that need to be considered when discussing
23 sustainability, elements that she characterised as 'threats' to the development of digital
24 projects, such as continuity of funding, data preservation (including choice of standards
25 and longevity of data) and the need for flexibility in order to change business models,
26 when necessary. More recently, libraries also have expressed a keen interest in
27 incorporating digital humanities centres and research activities as a way of ensuring
28 mutually beneficial development of methodological support, rather than promoting a
29 potentially problematic or unbalanced model service provision (see Montoya, 2016;
30 Posner, 2013; Siemens et. al. 2011). Needless to say, collaboration to support the
31 sustainability of project outputs features centrally in these discussions (see Irvine, 2015;
32 Green et. al., 2017; Vinopal and McCormick, 2013; Sweeney et. al., 2017; Montoya, 2016;
33 and Schaffner et.a;. 2014). A comprehensive review of literature published in LIS journals
34 about the topic of sustainability (Eschenfelder et. al., 2019) shows that this interest
35 remains strong, but also belies a number of the limitations on the understanding of a
36 digital humanities project and its contribution, discussed above. Sustainability was
37 generally understood as 'the continued operation of a collection, service, or organization
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 related to digital libraries, archives, and repositories over time' (Eschenfelder et. al.,
4 2019), but was also generally conceptualised as simply as 'a value (similar to
5 inclusiveness or efficiency) that ought to motivate action.' (Eschenfelder et. al., 2019).
6
7 This literature and others reviewed in the course of developing the CENDARI project
8 sustainability approach led to the conclusion that there were in fact a large number of
9 definitions of the term 'sustainability' in circulation, most of which envisioned the project or
10 its assets according to a different model or metaphor. While each of these models (which
11 will be discussed in the following sections) represents a frame of reference that
12 contributes greatly to the overall understanding of digital projects and their contribution,
13 the manner in which they seem to operate as silos may be a hindrance as well, losing
14 some of the richness and context that makes the digital project such a strong contributor
15 to contemporary humanities research.
16
17
18
19
20
21
22
23
24
25
26
27
28
29

30 MODEL 1: SUSTAINABILITY AS PRESERVATION OF DATA AND METADATA

31
32 It is notable in this respect that when it comes to discussing sustainability, research and
33 reflection coming out of the digital library community converges on a few key themes.
34 One of the most prominent of these perspectives views sustainability as a technical
35 function, related to the need to preserve digital objects, data, metadata and tools. For
36 example, an early document created within the University of North Texas Library
37 (Alemneh, Hastings and Hartman, 2002), indicates a series of steps to be taken for a
38 'healthy' data sustainability plan, including a Life Cycle Assessment of the Digital
39 Resources, Draft of a Metadata Architecture, Metadata Creation Workflow, and Metadata
40 Creation Tools. All of these elements are hugely useful contributors to the sustainability
41 of the work of the digital library, but are also indicative of a relatively narrow
42 conceptualisation of the impact of the digital library, viewing the data and collections of
43 the library as the sole object of a sustainability process. Other work reflects this same
44 conflation in different ways. The IFLA Manifesto for Digital Libraries (IFLA, 2010) for
45 example, also focusses on data as the key element to be sustained, but through the lens
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 of interoperability, rather than preservation. This theme continues to this day, with one
4 library-based author speaking of sustainability in the context of his work as applying to '...
5 a sub-genre of digital scholarship that primarily depends on the creation of relatively
6 static, combinatory content.' (Montoya, 2016, p. 606)
7
8
9
10

11 These examples of sustainability-related discussions reveal a relatively traditional
12 viewpoint, in which a digital project is viewed not so much as a new kind of scholarly
13 initiative, but an institution equivalent to its bricks and mortar precursor, built to house and
14 maintain collections, which does not need to prove and define its value on its own term,
15 but to manage threats to its fledgling existence.
16
17
18
19
20
21
22
23

24 MODEL 2: MODEL 3: SUSTAINABILITY AS MAINTENANCE OF AN ORGANISATION 25 OR INSTITUTION 26

27
28 Another conceptualisation of the digital project or infrastructure development that arises in
29 the literature around sustainability is of the project as a form of organisation or institution.
30 Often such approaches are focussed very strongly not so much on data as the key asset
31 to guarantee access to at a certain rate or level, but on creating a 'business model' for
32 maintaining the user-facing services and the organisation (including staff) created (see,
33 for example, Maron, Kirby Smith and Loy, 2014). The *Archives Portal Europe* project, for
34 example, decided to establish itself after a funded period of development as a
35 Foundation, as this organisational continuity was considered to be critical in order to
36 maintain the vital functions of the project and to further connect with other similar
37 initiatives. This creation of a Foundation was foreseen in APEnet project sustainability
38 report (Andr n et al., 2012), which described it as the best instrument by which to mitigate
39 the risks and ensure a future of sustainable growth: this mechanism was viewed as a way
40 to facilitate activities such as international cooperation and collaboration, sustainability of
41 the digital archival infrastructure and the assessment of use cases of digital services.
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57

58 This second mode of conceptualisation is also evident in a the research
59 undertaken by JISC and Ithaka S+R in 2008, which was published in the report
60

1
2
3 'Sustaining Digital Resources: an on-the-ground-view of Projects Today' (Maron and et
4 al., 2009), and also later into a 'Sustainability Implementation Toolkit' (Maron and Pickle,
5 2014). The analysis of the JISC/Ithaca report depicts a scenario where project
6 management strategies are crucial in the sustainability of a project after the end of its
7 funded period: dynamic and engaging leadership, identification of a value proposition,
8 creative cost management, and cultivation of diverse sources of revenue are all factors
9 that reflect the well-being of a digital project under this model.
10
11
12
13
14
15
16
17
18
19

20 MODEL 4: SUSTAINABILITY AS ENSURING TECHNICAL ROBUSTNESS FOR A 21 TOOL OR TECHNICAL PLATFORM 22 23

24 A further, more recently emerging model views the project as a *tool or technical platform*,
25 with the resulting sustainability proposal primarily taking into consideration issues and
26 practices such as migration and hosting of elements such as the repositories where the
27 data are stored and continued maintenance of work environments and specific tools. The
28 *TextGrid Project*, for example, has been hugely successful in ensuring that its platform
29 has remained accessible over a long period, continuing to make its services available to
30 users. At its best, this approach results in a broad focus on software durability,
31 documentation of processes and the modularity of services (Buddenbohn et al., 2015).
32
33
34
35
36
37
38
39
40

41 Similar to TextGrid, the DM2E (*Digital Manuscripts to Europeana*) project invested
42 significant effort in the development of a model to sustain its activities and tools after
43 project close. The goal of the project was to build tools and the communities to enable
44 humanities researchers to work with Manuscripts in the Linked Open Web. The project
45 worked on different levels of knowledge creation and transfer in order to enable scholars -
46 as well as third party projects such Europeana - to reuse the content they gathered and to
47 enrich it with annotation tools such as Pundit. In 2014 the project released a deliverable
48 on the sustainability of the project, entitled 'Result Transfer Plan' (Goldfarb et al., 2015).
49 This plan focussed largely on transferring technical management for the data, tools and
50 services developed by the project to the Europeana Digital Library to manage.
51
52
53
54
55
56
57
58
59
60

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 the 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

1
2
3 The reliance upon metaphors able to capture the essence of a digital project and align it
4 with a known field of objects and activities appears as a common thread in the writing and
5 thinking about sustainability. This alignment, once it is achieved, makes the digital
6 development less liminal, able to be imagined into the future as maintained at a certain
7 rate or level, and able to imply, by the nature of its alliance with collections, organisations,
8 software, brands or other model, a definition for what sustainability should mean.
9

10
11
12
13
14
15
16 These metaphors, however, become boundary objects in the sustainability
17 discourse, in two senses. First of all, they operate in the sense that Star and Griesmer
18 (1989) have described, as 'objects which are both plastic enough to adapt to local needs
19 and constraints of the several parties employing them, yet robust enough to maintain a
20 common identity across sites...weakly structured in common use, and become strongly
21 structured in individual-site use [with] different meanings in different social worlds but their
22 structure is common enough to more than one world to make them recognizable, a means
23 of translation' (393). By implying an identity that may or may not encompass their true
24 scale and scope, the metaphors that tend to underpin sustainability planning enable
25 understanding across intersecting social worlds (such as between researchers and
26 funders), but they may also betray the very developments they seek to make
27 comprehensible, or, indeed, to transcend the boundaries of. This is the second manner in
28 which the idea of the boundary object can be applied: literally, these conceptualisations
29 place boundaries upon the development, which may or may not encompass the full
30 richness of the projects and the elements to be taken into consideration in the
31 sustainability planning. The metaphors are, in fact, synecdoches, reflecting the project
32 team's desire to ensure their results are sustainable, but at the cost of reducing some of
33 the richness of the original development endeavor .
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52

53
54 A successful approach to sustainability for projects, and in particular for projects
55 claiming to develop research infrastructures, needs to be more comprehensive, therefore,
56 not just considering data or technology, community, communications or processes, but in
57 fact all of them simultaneously. In addition, it should focus on a project as a collection of
58
59
60

1
2
3 tangible and intangible assets, but also from the perspective of the potential user base for
4 these assets, and what these users consider valuable about them. Finally, it needs to
5 take its definition of sustainability from somewhere other than the norms of an imposed
6 model adapted via a metaphorical transfer of values and characteristics. Sustainability
7 must be underpinned not just by an understanding of 'what' and plan for 'how', but by
8 solid conceptions of 'why' and 'for whom.'

17 **The CENDARI Project Approach to Sustainability**

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

34
35
36
37
38
39
40
41
42
43
44
45
46
47
48 Another contributor to the early need to reconsider the proposed approach to
49 sustainability within CENDARI was the assumption that the project's relationship with a
50 durable European research infrastructure consortium, the DARIAH ERIC, would
51 guarantee that its results would be maintained. But a complex technical infrastructure
52 cannot simply be frozen in time and expected to continue to meet evolving needs. It
53 emerged early in the project that the communities that built CENDARI would need a more
54
55
56
57
58
59
60

1
2
3 active platform to ensure this, and the assets produced would need a broader base than
4
5 even DARIAH could provide in order to maximally benefit its potential users. It is from
6
7 these two realisations that the need to question and indeed redefine sustainability was
8
9 recognised within the project, and adopted as a piece of meta-research running in parallel
10
11 to development of the project's platform, data and research cases. Although both the
12
13 generic process and the specific actions resulting from this research match on some level
14
15 choices made for the CENDARI development, they also reflect the reality, identified by
16
17 Joris Van Zundert, of the 'fluidity' of research infrastructure, caught up in both the digital
18
19 information lifecycle and the creation of knowledge by end users, as well as the software
20
21 components (Van Zundert, 2012, 179).
22
23

24 The CENDARI sustainability plan therefore looked beyond the previous work
25
26 described above in a number of ways: first, it presented a view of *sustainability as a*
27
28 *process, rather than a state*, which begins with project conceptualisation and ends far
29
30 after project close; second, it views the *end goal of sustainability not as stasis, but as*
31
32 *transformation, and reuse*; finally, it views the project not as an organisation or technical
33
34 platform, but as a *collection of tangible and intangible assets* with potential value to other
35
36 users. As such, the plan proposed a comprehensive understanding of infrastructural
37
38 assets, and a model for sustainability that considered not only the most obvious end
39
40 product of the project (in CENDARI's case, a virtual research environment), nor even just
41
42 the project's tangible, reusable and customisable components - such as data and tools -
43
44 but also the project's intangible achievements, such as the know-how and the expertise
45
46 developed during its duration, and the community of users that had been built through the
47
48 project's testing and outreach phases.
49
50
51
52
53

54 **The CENDARI Assets and Multilevel Sustainability Action Plan**

55
56
57
58
59
60

1
2
3 The first component of the sustainability planning process, an audit of assets and reuse
4 scenarios, served to resist the common impetus to base sustainability measures on a
5 conceptualisation of the project that might not be comprehensive, agreed or indeed
6 explicitly recognised. This audit (which included both internal mapping and a stakeholder
7 validation meeting) refined project-level understanding of the full range of assets the
8 project had generated and how they could be maintained, shared and indeed passed on
9 to its key users for further development. This process identified seven categories of
10 assets as most likely to find future usage, each of which posed unique challenges in how
11 they could be captured, made visible and sustained. It was one of the greatest
12 challenges of the CENDARI sustainability planning process to ensure that for each of
13 these categories a solution to the challenges of sustainability was implemented, making
14 them findable and reusable in a contextualised manner, and preserving them in multiple
15 formats and multiple locations as needed. For each asset type, the audience for potential
16 future use would be different, and therefore the solutions proposed generally were as
17 well.

18
19
20 A brief overview of these seven categories follows. Although other projects would
21 almost certainly have a different mix of assets in their portfolio and place differing
22 emphasis on their importance, this list is indicative for the kinds of categories a project
23 should expect, and some of the measures that can sustain them.

24
25
26 The CENDARI **portal** was the most visible of its assets, representing the final
27 synthesis of the project's activities and its main point of access. For many projects, this
28 would be where sustainability planning would not only begin, but end. CENDARI
29 approached this sustainability challenge via a three pronged strategy, guaranteeing 3
30 years of access through the German arm of the DARIAH infrastructure but also ensuring
31 new communities and new approaches would be recruited to continue development. To
32 this end, the portal was also prepared and preserved as a virtual machine, known as
33 'CENDARI in a box,' which is available for research communities that would like to create
34 a new installation of the project environment with their own data or tools. In this way,
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 reuse of the technology underpinning the portal was separated from the portal itself, in
4 recognition that although its customised and dedicated ‘front door’ might seem the most
5 important element produced by any project, it is also generally the hardest to maintain,
6 and the most technically fragile over the medium to longer term.
7
8
9
10

11 The portal was not only useful in its complete final form, however, but also as a
12 collection of unique **services, tools and components** optimised to support digital
13 humanities research. This possible tactical reuse of the project outcomes was foreseen
14 from the beginning, and a very modular, service-oriented architecture was adopted for the
15 project. The tools therefore required a sustainable pathway outside of the portal as well,
16 although connecting tools with potential user bases is a constant challenge. The software
17 community practice of sharing software through GitHub, which Spiro and Smith estimate
18 to be used by around 20% of digital humanists (Spiro and Smith, 2016) was adopted, but
19 further awareness raising through dissemination and communication was also required to
20 ensure the maximal future use for the tools.
21
22
23
24
25
26
27
28
29
30
31
32

33 CENDARI holds a lot of **data** from different sources and in different formats, some
34 unique to the project, others well signposted elsewhere, all with different requirements
35 and expectations for sustainability. This has been its legacy as a project seeking to reuse
36 archival data for historical research, where the culture and ability to share data was
37 unevenly developed within the partner archives. The CENDARI data portal gives access
38 to this data, and the project’s data agreement and license were developed with DARIAH
39 as a co-signatory, ensuring that a permanent contact point for use and reuse questions
40 would be maintained. But DARIAH is not well-known as a data provider or source, and
41 this solution alone may not maximise visibility and reuse. Therefore a redeposit protocol
42 for unique data with an external trusted source was also facilitated. Institutions were also
43 offered the opportunity to reingest their own, transformed data, but in most cases the
44 institutions did not have an appropriate workflow in their side to make use of this, which
45 makes it all the more important that projects consider their data production as separate from
46 both the technology that might deliver it and the publications that might feature it. A
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 separate deposit of these assets as research, rather than cultural, data is therefore a
4 more sustainable approach to take under current conditions, where numbers of
5 repositories and finding aids for research data are growing and consolidating.
6
7

8
9 The CENDARI **Archival Research Guides**, which combined the elements of an
10 archival finding aid with a scholarly interpretation in a dynamic, multimedia format, existed
11 as a particular subset of the data held unique to CENDARI, but their status as both
12 primary and secondary research sources justified their consideration as an asset class in
13 themselves, in particular because of the manner in which they challenged existing norms
14 of publication, communication and evaluation in the discipline of history. As extended and
15 enhanced publications, incorporating analysis, links to data sources, multimedia objects,
16 and links to project ontologies, these guides needed to be delivered within the project
17 portal to be used to their full potential. But to sustain these unique works of scholarship
18 only in that format would again potentially limit their visibility. They are therefore offered in
19 several export formats, as well as becoming the focus of both a review publication and a
20 research paper in a mainstream (not digital) historical journal. In this way, their
21 contributions to scholarship can be recognised as independent from the format in which
22 they have been delivered, while the affordances of that format can still be harnessed.
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38

39 Given how particular many of the project experiences in building for the DH
40 community had been, a specific audit of CENDARI's **tacit knowledge** was also
41 undertaken, and several white papers and process oriented toolkits emerged from the
42 project on the foot of this (including, for example, the 'White Book of Archives,' (Beneš et
43 al., 2015) documenting the project's experience of federating highly heterogenous data
44 from traditional collection holding institutions).
45
46
47
48
49
50

51 Perhaps the least easily defined and sustained aspects of the CENDARI project
52 were the **communities** - mixed and homogenous groups of historians and other
53 humanistic scholars, collections experts and technologists - it brought together and
54 fostered for four years. Interconnectivity between cognate projects was a key resource for
55 facilitating this, as some communities will have interests across these projects, and
56
57
58
59
60

1
2
3 networks can and should be shared. Some voices in infrastructure studies do recognise
4 the importance of considering this layer, encouraging projects to view ‘technical
5 infrastructure as a human and community asset in need of maintenance and support,
6 rather than a technical artefact in need of service management’ (Smithies et.al., 2019).
7
8 But adopting this perspective that puts the community first remains a challenge, which
9 many projects face with the development of more technical infrastructure, such as the
10 Modernist Commons focus on sustaining its community via ‘Web-accessible repository of
11 digitized resources, a Web-based workbench of digital editing tools, and a customizable
12 publication interface’ (Irvine, 2015, p. 5). The CENDARI team felt that such an approach
13 would not be appropriate, however, given the place of the digital in the workflows of the
14 community it served, and the nature of the interactions and consensus required to ensure
15 true sustainability. For this reason, the project decided to use another DARIAH
16 mechanism, the Working Group, to provide a structure for continued development of
17 project concerns and assets, monitoring the value it is providing at regular time steps and
18 taking action to maintain or indeed increase it.
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36

37 **High-Level Principles to Foster Sustainability**

38
39
40
41 In analysing the full possible range of reuse scenarios for these classes of assets, the
42 CENDARI project developed a set of five high-level principles to secure the reuse
43 potential of any digital project. These are:
44
45
46

47 1. EMBED INTO PARTNERSHIPS and NETWORKS: Users will need to know
48 how any new tool or service fits in with the existing landscape they know. Working with
49 other projects and organisations will help you to build your networks as well as your
50 credibility, as well as giving you possible partners at the end of the project to take on and
51 further development of your data sets or tools. By sharing widely and embedding in
52 networks, you can build a solid basis for a core group to continue oversight of the project
53
54
55
56
57
58
59
60

1
2
3 results: this is easier to facilitate if a new umbrella identity can be established. The DM2E
4 transfer of results to Europeana is an excellent example of how this principle can be
5 delivered upon.
6
7
8

9 2. DEVELOP WITH MODULARITY AND OPENNESS: Designing a digital project
10 in a way that allows single elements to be reused and implemented by other research
11 infrastructures and projects will add a rich further layer to the possible scenarios for their
12 reuse. In addition, projects should use common platforms, such as GitHub for software or
13 an open repository for reports, and known and accepted standards in their work, as
14 aspect also recognised in other projects (Vinopal and McCormick, 2013, p. 31). These
15 may be drawn from a number of relevant communities (TEI for scholarly editions, EAD or
16 ISAD for archival material, etc.) but maintaining reference to a larger context for the
17 organisation of data should be viewed as an essential element of its potential reusability.
18
19
20
21
22
23
24
25
26
27

28 3. DOCUMENT, VALIDATE and SHARE: Knowledge sharing across the project
29 and beyond should be not just the norm, but a process projects actively consider and
30 invest in. This should be implemented not just at the level of project heads, but across
31 the cognate skills areas within and between projects, with communications people
32 building relationships with other communications people, for example. Mapping project
33 structures for this purpose against each other should occur as a matter of course. Project
34 publications should cover both scientific and management aspects of the project, with
35 different audiences in mind for different pieces, and, where project publications challenge
36 community norms, a source of external validation should be identified for them via open or
37 tailored peer review, or consultation mechanisms. To be successful, this tenet should be
38 interpreted as broadly as possible: the CENDARI project was in the very fortunate position
39 to have a preferential relationship with a durable, transnational research infrastructure, but
40 a committed institution, network, or indeed a commercial publisher may also play a
41 significant role for a project, as in Laura Mandell's experience partnering with Gale for the
42 improvement of the EEBO transcriptions: when asked 'What is your sustainability plan?'
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 for the grant deliverables you are producing for the eMOP project, I would answer, "Gale.
4 Gale is my sustainability plan." (Mandell and Grumbach, 2015, p.3)
5
6

7 In addition, every project should have a thorough data management plan,
8 explaining what data it will collect or create data and where it will be left, including as well
9 any services the project will create to manage this data and how they will be sustained.
10
11 Technical documentation of the tools and their integration should be openly available.
12
13 Finally, a 'Tacit Knowledge Audit' should be built into the project so that applied
14 knowledge and lived experience does not get lost with project closure, but can rather be
15 shared explicitly in an appropriate format. In particular assets such as usage policies,
16 management plans, cooperation agreements and the like can be of great value to
17 successor projects.
18
19
20
21
22
23
24
25

26 4. MAINTAIN CLARITY OF PURPOSE and SIMPLICITY OF ACCESS: This can
27 be facilitated by designing, for example, a data ingestion cycle that is capable of being
28 rolled out as an easily managed service, providing end users with simple instruments or
29 forms to contact the project's team, to add content, report bugs and query usage of tools.
30
31 Maintaining a consistent central communication point, even after active project close, and
32 ensuring that research work can be accessed reliably in a variety of easy to find, relevant
33 formats and locations (such as a trusted digital repository or traditional journal) will
34 support this objective as well. Finally, a project will benefit from being clear about what it
35 has created or collected, what its value is, and for whom.
36
37
38
39
40
41
42
43
44

45 5. APPROACH SUSTAINABILITY AS A PROCESS, NOT AN END PRODUCT.
46 The CENDARI sustainability planning process was comprised of a series of stages,
47 stretching from pre-planning and mature project to closure and post-project actions, each
48 of which contributed to the overall, holistic sustainability strategy. This cycle was intended
49 to counteract the natural impetus within projects to view sustainability as a concern only
50 for the final phase of the project, rather than one to be integrated into the project's
51 development and even its conception. Implementing such a process requires focus and
52 resources, but results in a far better and more grounded hand-over phase for the project
53
54
55
56
57
58
59
60

1
2
3 from a funded period to a post-funded period. To capture the benefits of this process,
4 project teams should make sustainability a part of their initial proposals, including at a
5 minimum a sustainability meeting at the project mid-point and a sustainability report as a
6 project outcome
7
8
9
10

11 From a technical standpoint, projects should seek to guarantee the availability of
12 their basic technical formats for 3 to 5 years after close. This may seem short, but it is not
13 necessarily reasonable to expect or support a longer window, as beyond that time it would
14 need to be significantly updated anyway. Committing to a defined, shorter span will
15 encourage early and sustained dialogue between stakeholders. Where projects continue
16 to prove their utility to new and old user communities within and beyond this basic
17 window, successor funding projects to focus on new user groups or new functionalities
18 should naturally come to be sought, ensuring the project maintains the ability to be
19 'evolving and involving.' Where the usage becomes more sporadic and specialised over
20 time, an alternative should be sought, such as clearly and richly documenting the user
21 interface and tools, such as search queries, but ultimately extracting the data layer from
22 the software designed to deliver it for deposit in a long term, stable, open resource, a
23 process Geoffrey Rockwell and his coauthors refer to as 'Burying a Dead Project'
24 (Rockwell et al., 2014). The model offered by the Kings Digital Lab to sustain digital
25 projects is groundbreaking in this way, offering as it does a choice of hosting options
26 under a service level agreement, via migration to another institution, or as archived data
27 (Smithies et. al, 2019). Even under such an agreement, however, a project should be
28 sure to think about what might happen in the future, for example when a new host, in
29 particular a commercial one, looks to migrate projects. As Laura Mandell reflected in the
30 context of her work with Gale, 'We do need to have in our contracts with such vendors
31 that,when the company or its successors no longer wishes to sell and sustain a digital
32 catalogue, it must pass it to libraries for dark archiving.' (Mandell and Grumbach, 2015,
33 p3)
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

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

1
2
3 research that is collaborative, digitally-enabled, open-ended, and performed 'at scale'
4
5 outside of traditional institutions and networks can be designed, developed, and validated
6
7 holistically. We hope the experiences of the CENDARI project can make an ongoing
8
9 contribution to this cultural change within our communities and disciplines.
10
11
12

13 **Funding**

14
15 This work was supported by the European Commission's Seventh Framework
16
17 Programme [grant number 284432].
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

References

- 1
2
3
4
5
6
7 Alemneh, D., et. al. (2002), "A Metadata Approach to Preservation of Digital Resources",
8
9 *First Monday* 7(8).
10
- 11 Andrén, P., et. al. (2012), *APEnet Sustainability Report*
12
13 [http://pro.europeana.eu/files/Europeana_Professional/Projects/Project_list/APEnet/Deli](http://pro.europeana.eu/files/Europeana_Professional/Projects/Project_list/APEnet/Deliverables/D5.4%20Sustainability_Report.pdf)
14
15 [verables/D5.4%20Sustainability_Report.pdf](http://pro.europeana.eu/files/Europeana_Professional/Projects/Project_list/APEnet/Deliverables/D5.4%20Sustainability_Report.pdf) (accessed 24 July 2017).
16
17
- 18 APE Foundation. <http://www.archivesportaleuropefoundation.eu/> (accessed 24 July
19
20 2017).
21
- 22 Beneš, J., et. al. (2015), *The CENDARI White Book of Archives*,
23
24 <http://www.cendari.eu/sites/default/files/WhiteBook-Web.pdf> (accessed 24 July 2017).
25
26
- 27 Buddenbohm, S., et. al. (2015), "Success Criteria for the Development and Sustainable
28
29 Operation of Virtual Research Environments", *D-Lib Magazine*, 21(9/10).
30
- 31 CENDARI. <http://www.cendari.eu> (accessed 24 July 2017).
32
- 33 Dombrowski, Q. (2014), 'What Ever Happened to Project Bamboo?' *Literary and*
34
35 *Linguistic Computing*, 29(3): pp. 326–339.
36
- 37 Eschenfelder, Kristin R., et. al. (2019), "A Nine Dimensional Framework for Digital Cultural
38
39 Heritage Organizational Sustainability", *Online Information Review*, 43.2 (2019), pp.
40
41 182–96 <<https://doi.org/10.1108/OIR-11-2017-0318>>
42
- 43 Goldfarb, D., et. al. (2014), *DM2E Results Transfer Plan*,
44
45 http://dm2e.eu/files/D5.10_1.0_Results_Transfer_Plan_140730.pdf (accessed 24 July
46
47 2017).
48
- 49 Green, H. et. al. (2017), *Scholarly Commons Digital Humanities Needs Assessment Study*
50
51 (University of Illinois, 2017)
52
53 <[https://www.ideals.illinois.edu/bitstream/handle/2142/100081/Final_DHNeedsAssess](https://www.ideals.illinois.edu/bitstream/handle/2142/100081/Final_DHNeedsAssessment_Report.pdf?sequence=2&isAllowed=y)
54
55 [ment_Report.pdf?sequence=2&isAllowed=y](https://www.ideals.illinois.edu/bitstream/handle/2142/100081/Final_DHNeedsAssessment_Report.pdf?sequence=2&isAllowed=y)>
56
57
- 58 IFLA Governing Board (2010). *IFLA/UNESCO Manifesto for Digital Libraries*,
59
60 <http://www.ifla.org/digital-libraries/manifesto> (accessed 24 July 2017).

- 1
2
3 Irvine, D. (2015), "From Angel to Agile: The Business of the Digital Humanities", *Scholarly*
4 *and Research Communication*, 6.2, pp. 1-8.
- 5
6
7 Maron, N., et. al. (2009), *Sustaining Digital Resources: An On-the-Ground View of*
8 *Projects Today*, JISC and Ithaka S+R, [http://www.sr.ithaka.org/wp-](http://www.sr.ithaka.org/wp-content/mig/reports/4.17.2.pdf)
9 [content/mig/reports/4.17.2.pdf](http://www.sr.ithaka.org/wp-content/mig/reports/4.17.2.pdf) (accessed 24 July 2017).
- 10
11
12
13 Maron, N., Pickle, S. (2014). *Sustaining the Digital Humanities: Host Institution Support*
14 *Beyond the Start-up Phase*, Ithaka S+R,
15 <http://www.sr.ithaka.org/publications/sustaining-the-digital-humanities/> (accessed 24
16
17
18
19
20
21 July 2017).
- 22 Maron, N., Pickle, S., (2014) *Sustainability Implementation Toolkit: Developing an*
23 *Institutional Strategy for Supporting Digital Humanities Resources*, Ithaka S+R,
24 <http://www.sr.ithaka.org/publications/sustainability-implementation-toolkit/> (accessed
25
26
27
28
29
30
31
32
33
34 24 July 2017).
- 35 Mandell L. and Grumbach, E. (2015), "The Business of Digital Humanities: Capitalism and
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60 Enlightenment", *Scholarly and Research Communication*, 6.4 (2015), 1–9
- Montoya, R. (2016), "Advocating for Sustainability: Scaling-Down Library Digital
Infrastructure", *Journal of Library Administration: Vol 56, No 5'* pp. 503-620.
<https://www.tandfonline.com/doi/abs/10.1080/01930826.2016.1186969?scroll=top&neededAccess=true&journalCode=wjla20> [accessed 13 November 2019]
- Posner, M. (2013) "No half measures: Overcoming common challenges to doing digital
humanities in the library", *Journal of Library Administration*, 53(1), pp. 43–
52. <http://doi.org/10.1080/01930826.2013.756694>
- The Pundit. <http://thepund.it/> (accessed 24 July 2017).
- Rockwell, G. et.al. (2014), "Burying Dead Projects: Depositing the Globalization
Compendium", *Digital Humanities Quarterly*, 8.2.
- Schaffner, J. et. al. (2014), *Does Every Research Library Need a Digital Humanities*
Center?, [http://oclc.org/content/dam/research/publications/library/2014/oclcresearch-](http://oclc.org/content/dam/research/publications/library/2014/oclcresearch-digital-humanities-center-2014.pdf)
[digital-humanities-center-2014.pdf](http://oclc.org/content/dam/research/publications/library/2014/oclcresearch-digital-humanities-center-2014.pdf)> (accessed 8 November 2019)

- 1
2
3 Siemens, L. et. al. (2011), "A tale of two cities: implications of the similarities and
4 differences in collaborative approaches within the digital libraries and digital humanities
5 communities", *Literary and Linguistic Computing*, Vol. 26, No. 3, pp. 335-348.
6
7
8
9
10 Smithies, J. et.al. (2019), "Managing 100 Digital Humanities Projects: Digital Scholarship
11 & Archiving in King's Digital Lab", *Digital Humanities Quarterly* 13.1.
12
13
14 Spiro, L. and Smith, S.M. (2016) "Evaluating GitHub as a Platform of Knowledge for the
15 Humanities" paper presented at the Digital Humanities Conference, 11-14 July,
16 Krahow, Poland, available at: [http://dbs.uni-](http://dbs.uni-leipzig.de/file/aumueller05wiksar.pdf)
17 [leipzig.de/file/aumueller05wiksar.pdf](http://dbs.uni-leipzig.de/file/aumueller05wiksar.pdf) (accessed 9 November 2019).
18
19
20
21
22 Star, S.L. and Griesemer, J. (1989). "Institutional Ecology, 'Translations' and
23 Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate
24 Zoology 1907-39", *Social Studies of Science*, 19(3): pp. 387-420.
25
26
27
28 Sweeney S.J., Flanders, J., and Levesque, A. (2017), "Community-Enhanced Repository
29 for Engaged Scholarship: A Case Study on Supporting Digital Humanities Research",
30 *College & Undergraduate Libraries*, 24.2-4 (2017), pp. 322-36
31
32
33
34 TextGrid. <https://textgrid.de/en> (accessed 24 July 2017).
35
36
37 Van Zundert, J. (2012). "If you build it, will we come? Large scale digital Infrastructures as
38 a dead end for digital humanities", *Historical Social Research Journal*, 37(3): pp. 165-
39 186.
40
41
42
43 Vinopal, J., and McCormick, M. (2013), "Supporting Digital Scholarship in Research
44 Libraries: Scalability and Sustainability", *Journal of Library Administration*, 53 (2013),
45 pp. 27-42
46
47
48
49 Warwick, C., et. al. (2008), "If You Build It Will They Come? The LAIRAH Study:
50 Quantifying the Use of Online Resources in the Arts and Humanities through Statistical
51 Analysis of User Log Data", *Literary and Linguistic Computing*, 23(1): pp. 85-102.
52
53
54
55
56 Zorich, M. (2003). *A Survey of Digital Cultural Heritage Initiatives and Their Sustainability*
57 *Concerns*, <https://www.clir.org/pubs/reports/pub118/pub118.pdf> (accessed 24 July
58 2017).
59
60