



# THE UNIVERSITY *of* EDINBURGH

## Edinburgh Research Explorer

### Digital editions of text

**Citation for published version:**

Franzini, G, Terras, M & Mahony, S 2019, 'Digital editions of text: Surveying user requirements in the Digital Humanities', *Journal on Computing and Cultural Heritage*, vol. 12, no. 1. <https://doi.org/10.1145/3230671>

**Digital Object Identifier (DOI):**

[10.1145/3230671](https://doi.org/10.1145/3230671)

**Link:**

[Link to publication record in Edinburgh Research Explorer](#)

**Document Version:**

Peer reviewed version

**Published In:**

Journal on Computing and Cultural Heritage

**Publisher Rights Statement:**

© ACM, 2019. This is the author's version of the work. It is posted here by permission of ACM for your personal use. Not for redistribution. The definitive version was published in Journal on Computing and Cultural Heritage ,VOL 12, ISS 1, February <https://dl.acm.org/citation.cfm?doid=3313804.3230671>

**General rights**

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



# Digital Editions of Text: Surveying User Requirements in the Digital Humanities

Greta Franzini<sup>a</sup>, Melissa Terras<sup>b</sup>, Simon Mahony<sup>a</sup>

<sup>a</sup>University College London, UCL Centre for Digital Humanities

<sup>b</sup>University of Edinburgh, College of Arts, Humanities and Social Sciences

## Abstract

This paper presents the findings of a web survey designed to better understand the expectations and use of digital editions of texts. The survey, modelled upon a detailed analysis of 242 projects, recorded 218 complete responses, shedding light on user requirements of digital editions. Specifically, the survey indicates that issues of data reuse, licensing, image availability, and comprehensive documentation are the most requested features of digital editions although ones which seldom are provided. This analysis feeds into previous studies on good practice in building Digital Humanities resources and puts forward practical recommendations for both creators and funders of digital editions in an effort to promote a stronger consideration of user needs. This survey will be of interest to those who produce digital editions of texts, including developers and engineers, and will also be of interest to those who commission and fund these projects, such as universities, libraries and archives, whose documentary collections are often showcased in digital editions.

## Introduction

*"good idea this survey!"<sup>1</sup>*

Since the early 1980s there has been a great amount of effort put into the creation of digital textual editions. A digital textual edition can be broadly defined as a reconstruction of a literary text inserted in the text's print or manuscript tradition and which makes use of digital technologies to reproduce any level of detail of that text. To better understand the methods, technologies and rationale behind the production of digital editions, we created a Catalogue that has so far identified and categorised 256 of these projects produced worldwide.<sup>2</sup> The user-side of digital editions, however, has not received equal attention. The study described in this article follows on from the few existing efforts in this relatively unexplored area and sought to answer the following research question: *what are the expectations of digital editions in the Digital Humanities and how do these correlate with existing digital editions?*

Here, we compare the findings of our analysis of digital editions, which detailed their attributes and implementation, with the results of a widespread user survey that identified the important and salient features the community most desires. We compare and contrast the two, resulting in a series of practical recommendations for creators and funders of digital editions, indicating that issues of data reuse, licensing, image availability, and comprehensive documentation are the most requested features of digital editions (although ones which seldom are provided). This

---

<sup>1</sup> Comment left by a participant of this survey (in reply to Question 20 - see further below).

<sup>2</sup> See: <https://dig-ed-cat.acdh.oew.ac.at>. This analysis was undertaken by the first author. All numerical considerations made in this paper with regard to the Catalogue of Digital Editions reflect the state of the project as of May 2017. Since then, new digital editions have been added and these are not included in the data analysis described here.

paper is the first comparative analysis of the differences between what the Digital Humanities community who builds online digital editions of texts provides, and what the user-base wants. It therefore provides useful guidance for those producing digital editions of texts and will also be of interest to those who commission, fund, and support these projects, including universities, libraries and archives, whose documentary collections are often showcased in digital editions.

## Digital Editions

In the Digital Humanities, the term 'digital edition' is typically used to denote projects that make use of digital technologies to reproduce editions and transcriptions of literary texts, be those inscribed on tablets or penned on papyrus, vellum or paper. Digital editions themselves are to be considered an asset of cultural heritage, as Tomasi explains:

Le edizioni digitali fanno parte del patrimonio culturale e vanno quindi valorizzate al pari delle raccolte librerie, archivistiche e museali, anche in considerazione della realizzazione di digital libraries nella forma di aggregatori di risorse come strumento di accesso integrato al patrimonio culturale [...].

(2013, p. 25)<sup>3</sup>

The multiple definitions and types of (digital) edition that exist are listed in the *Lexicon of Scholarly Editing*.<sup>4</sup> Scholars distinguish between *digital* and *digitised* editions –the former indicating interactivity as opposed to mere reproductions of analogue material–, or between *digital editions* and *digital scholarly editions* –the latter describing a project with a strong critical component.

Digital editions are a key interest in the Digital Humanities. For almost forty years, the community has been building these resources (Vanhoutte 2010),<sup>5</sup> which now number in the hundreds (Franzini et al., 2016). A handful of digital editions are worth mentioning to briefly show the range of contributions they make to textual scholarship. One of the earliest digital editions produced was *The Canterbury Tales Project*; conceived in 1989, published seven years later on CD-ROM (Robinson, 2006) and now part of a collaborative editing platform<sup>6</sup>, this project bears witness to developments over the past twenty years. Another example is the *Codex Sinaiticus* project, widely recognised as one of the first endeavours to virtually reunite a disbanded work, in this case a fourth century manuscript of the Christian Bible whose leaves are scattered between the UK, Russia, Egypt and Germany.<sup>7</sup> The *Briefwechsel Sauer-Seuffert* project builds a body of correspondence into a dynamic timeline.<sup>8</sup> And, finally, the *Digital Dead Sea Scrolls* project mimics the act of rolling through parchment scrolls allowing users to interactively select regions on an image to inspect the corresponding text.<sup>9</sup> These examples show the variety of projects available, often carried out in order to provide wide access to cultural heritage, which would otherwise be restricted to experts (e.g. researchers, conservators

---

<sup>3</sup> The following is an English translation of the quote by the first author: "Digital editions form part of cultural heritage and should, therefore, be valued equally to library, archival and museum collections; they should also be considered for inclusion in digital libraries –intended here as resource-aggregators– as access tools to cultural heritage [...]."

<sup>4</sup> See:

<https://web.archive.org/web/20170907122807/http://uahost.uantwerpen.be/lse/index.php/lexicon/edition-digital/>

<sup>5</sup> Van Hulle has published an annotated list of key theoretical works in the field (2006), to which the authors would add Hockey (2000) and more recent volumes, including Sutherland and Deegan (2012), Pierazzo (2016), Driscoll and Pierazzo (2016), Sahle (2014) and McGann (2014).

<sup>6</sup> Available at: <http://www.textualcommunities.usask.ca/web/canterbury-tales>

<sup>7</sup> The *Codex Sinaiticus* project is available at:

<https://web.archive.org/web/20170907124405/http://www.codexsinaiticus.org/en/>

<sup>8</sup> See: <https://web.archive.org/web/20170907124016/http://sauer-seuffert.onb.ac.at>

<sup>9</sup> See: <https://web.archive.org/web/20170907124248/http://dss.collections.imj.org.il>

and other individuals who are specifically granted access) or to special occasions (e.g. temporary exhibitions in libraries or museums).

## Related Work

Despite the large number of digital editions being created, no extensive study has been conducted to discuss the needs of *users* of digital editions, with the consequent result that anybody attempting to do such research, or build a digital edition, does so with very little prior evidence as to usage patterns or requirements. In 2008, the LAIRAH (Log analysis of Internet Resources in the Arts and the Humanities) initiative<sup>10</sup> analysed twenty-one digital resources and found that:

P[rincipal]I[nvestigator]s [...] infer[red] user requirements from their own behaviour.

(Warwick et al., 2008, p. 390)

In 2010, Terras stated that project reports typically disclose information about the number of visits, downloads and time spent on a particular page but not the users' actual experiences, and that few users self-report (Terras 2010, pp. 6-7). In the same year, an article entitled 'Electronic Editions for Everyone' hinted at users but, in fact, mainly discusses the difficulties creators face in building digital editions for a wide range of users, *not* the identity of these users nor their needs for digital editions (Robinson, 2010). In 2012, Hughes' volume *Evaluating and Measuring the Value, Use and Impact of Digital Collections* examined, as the title suggests, approaches to evaluating digital *collections*, not digital editions, even though digital editions can, in certain cases, be considered collections or archives in the sense that they collect documents that make-up a particular literary work:

In a digital environment, *archive* has gradually come to mean a purposeful collection of surrogates. As we know, meanings change over time, and *archive* in a digital context has come to suggest something that blends features of editing and archiving. To meld features of both — to have the care of treatment and annotation of an edition and the inclusiveness of an archive — is one of the tendencies of recent work in electronic editing. One such project, the *William Blake Archive*, was awarded a prize from the Modern Language Association recently as a distinguished scholarly edition. (Price, 2009)

In 2013, Dot Porter published an article entitled 'Medievalists and the Scholarly Digital Edition' in which she summarised the results of two surveys she conducted in 2002 and 2011 to learn more about *medievalists'* attitudes toward electronic resources, including digital editions (Porter, 2013). Both surveys asked participants from different departments to indicate their use of print and digital editions in a five-point Likert scale (*Electronic only; Electronic mostly; Electronic and print; Print mostly; Print only*).

Her first survey, circulated in 2002, addressed a random controlled sample of 92 faculty members of medieval studies in different departments and was delivered via mail (86 participants) and email (6 participants); of the 92 surveys sent, 43 (46.7%) were completed (ibid., p. 5). A second survey was prepared in 2011 and targeted a wider audience. To do so, Porter again selected a random controlled sample of faculty members but also shared the survey via Twitter, Facebook and mailing lists. The total number of respondents of the 2011 survey was 169 (27 from faculty, 142 from the open survey). The difference between the two surveys,

---

<sup>10</sup> See: <https://web.archive.org/web/20171008085208/http://www.ucl.ac.uk/infostudies/LAIRAH/>

Porter informs, was that the 2002 survey did not ask about the use of electronic books (i.e. those available on e-readers such as Kindle, Nook and the iPad), which were only popularised a decade later (ibid., p. 6). The results of the 2002 survey indicated a preference of print over electronic editions, possibly because of the lower number of electronic resources available at that time. In the 2011 survey, a majority of all groups except for music and history again reported a preference for print editions. In summary, Porter's experiments showed that the nine-year gap between surveys did not record a large shift in medievalists' usage of digital editions, a behaviour which, she posits, might be explained by their lack of interest in this type of resource (as opposed to electronic journals and facsimiles) and "a lack of understanding by non-digital-editing medievalists about what exactly a digital scholarly edition is" (ibid., p. 14).

More recent monographs on digital scholarly editing, such as *Digital Critical Editions* by Apollon et al. (2014) also do not give much space to user-studies (Franzini, 2015).

A 2015 MIT online announcement searching for people to test the *Infinite Ulysses* digital edition is one of a very small number of initiatives interested in understanding how a project is used (Visconti, 2015). More worryingly, at the 2015 conference of the *European Society for Textual Scholarship* (ESTS) entitled 'Users of Digital Editions'<sup>11</sup> –the first effort of a well-established scholarly network to examine users and their needs– only one paper touched upon the topic and that described a *user interface*.<sup>12</sup> Interfaces of digital editions were also the central theme of the 2016 conference *Digital Scholarly Editions as Interfaces*<sup>13</sup> organised by the Digital Scholarly Editions Initial Training Network (DiXiT)<sup>14</sup>. In the same year, Dillen and Nyet (2016) introduced three categories of users modelled upon three types of use the authors envisage for digital editions, confirming that users and user needs are not well known but often *inferred* by creators:

At the most basic level of interest, users are looking for simple browsing functionalities. To satisfy these users, editors will want to present the materials within an attractive and intuitive interface. At a more advanced level of interest, users will want to research the materials the DSE [Digital Scholarly Edition] has to offer, and access them in non-linear ways. To reach those users, editors will need to provide indexes, advanced search options, advanced textual comparison options, to open the corpus up for analysis in a standardized format, etc. Finally, at the highest level of interest, there are meta-users, who want to re-use the DSE's data for their own purposes: to write their own transcriptions of the DSE's facsimiles (and publish the results), to build their own interface around the data

---

<sup>11</sup> The flyer of the ESTS 2015 conference can be accessed here: <http://cts.dmu.ac.uk/ESTS/ESTS-flyer.pdf>. The ESTS homepage is available at:

[https://web.archive.org/save/\\_embed/https://textualscholarship.eu/](https://web.archive.org/save/_embed/https://textualscholarship.eu/)

<sup>12</sup> The talk, entitled "Beyond Google Search: Editions as Dynamic Sites of Interaction", was given by PhD student Shane McGarry from Maynooth University. McGarry reported on the user survey undertaken for *The Woodman Diary* project to evaluate the usability of the digital edition's *user-interface*. The project is available at:

<https://web.archive.org/web/20170907125220/http://dhprojects.maynoothuniversity.ie/woodman/>

<sup>13</sup> More information about the conference is available at:

[https://web.archive.org/save/\\_embed/https://informationsmodellierung.uni-graz.at/en/events/archive/digital-scholarly-editions-as-interfaces/](https://web.archive.org/save/_embed/https://informationsmodellierung.uni-graz.at/en/events/archive/digital-scholarly-editions-as-interfaces/). Video recordings of the talks given

at this conference are available at the ZIM ACDH YouTube channel:

[https://web.archive.org/channel/UCFb\\_IysRdxHsvS9dZw1O1lw](https://web.archive.org/channel/UCFb_IysRdxHsvS9dZw1O1lw)

<sup>14</sup> For more information about DiXiT, see:

<https://web.archive.org/web/20170907125622/http://dixit.uni-koeln.de/>

the DSE provides, or to perform functionalities the DSE does not (yet) offer (and publish the results).

(Dillen and Neyt, 2016, p. 3)

The closest we have come to exploring the gap between creators and users of digital editions is the relatively recent establishment of the *RIDE Review Journal of Digital Scholarly Editions and Resources*, which excels at evaluating digital scholarly editions through the publication of individual reviews<sup>15</sup>: however, no overarching analysis has been done on community investment in digital resources, as compared to user desires of digital resources. Users of digital editions should be able to share their concerns, needs and feedback with regard to the creation of these projects. This large gap should no longer be ignored as excluding users' voices from the early development stages of digital edition can lead to neglect:

In the case of digital humanities large amounts of public funding is wasted if a resource is not used.

(Warwick et al., 2006, p. 16)

Again, in reference to Digital Humanities resources in general, not digital editions specifically:

Very few projects maintained contact with their users or undertook any organised user testing, and many did not have a clear idea how popular the resource was or what users were doing with it.

(Warwick et al., 2007, p. 2)

These studies confirm that by involving users in the creation process of a digital edition the risk of neglect might be reduced and, therefore, the chances of it being used in the long-term increase. So why are creators not engaging potential users more actively? Aside from the findings of the aforementioned LAIRAH study, the most recent article on the topic reports that usability testing in Digital Humanities is not yet widely established for a variety of reasons, including the absence of usability tests or survey templates in the Humanities for specific services and the vague or unspecific research questions feeding the development of the services (Bulatovic et al., 2016):

The aim of those of us designing resources in digital humanities, therefore, remains analogous to this. We must understand the needs and behaviours of users. As a result of this understanding, we must design resources that fit well with what our users already do, while providing advantages in terms of convenience, speed of access, storage capacity and innovative information tools that digital publication affords. If we do so, there is every chance that such resources will be used and will help to make possible new kinds of scholarship that would be inconceivable without digital content, tools and delivery mechanisms.

(Warwick, 2012, p. 19)

This article reports on the first initiative to bring together a *varied* group of users of digital editions with a view to identifying needs and expectations. Before turning to this study, however, it is first necessary to introduce the *Catalogue of Digital Editions* as the data contained therein will help us better evaluate user needs against the digital (scholarly) editions being built in the Digital Humanities.

---

<sup>15</sup> Available at: <https://web.archive.org/web/20170907125810/http://ride.i-d-e.de/>

## The Catalogue of Digital Editions

The Catalogue of Digital Editions was first launched by the first author in 2012 in response to a need to survey and identify best practice in the field of digital editing. Inspired by Patrick Sahle's *Catalog of Digital Scholarly Editions*,<sup>16</sup> the first resource of this kind, this Catalogue sought to examine online digital editions in order to extract statistical information and thus better observe developments in the field.

Franzini et al. (2016) described the early days of the Catalogue. In summary, data collection began as a spreadsheet for the first author's personal use; there, she recorded a small number of identifying features for every project, including name, URL, name of manager(s), responsible institution(s), and the historical period to which the edited text belonged. In time, with the addition of more projects the number of cataloguing criteria also increased to address as many dimensions as possible (e.g. type of licence, availability of images, search functionality, provision of indices and download options, etc.). Once a stable table structure was reached, a decision was made to publicly share the spreadsheet as a collaborative document for others to use and edit should they so wish. Although the Catalogue did not see a large number of external contributions, the interest shown over a period of two years prompted the first author to maintain this list as an open and citeable data resource in GitHub.<sup>17</sup>

In the summer of 2016, the Catalogue of Digital Editions turned into a collaboration with the Austrian Centre for Digital Humanities at the Austrian Academy of Sciences; the goal of the collaboration is to give the Catalogue dynamic functionality, shaping it into a web application and dissemination platform to allow users to contribute, download, browse, search, visualise and filter data around their research interests. In April 2017, the German *Datenbank-Infosystem* (DBIS) added the Catalogue to its list of open scientific databases as a service for use in over 300 libraries in Germany.<sup>18</sup> As of September 2017, the number of cataloguing features used by the Catalogue is 49, and users have requested an additional five features to be added.<sup>19</sup>

Ultimately, as outlined by the digital publications manager of the J. Paul Getty Trust, the Catalogue proposes an evaluation model for digital editions and similar digital educational resources:

Though limited to its own particular subset of digital publishing activity, Franzini's Catalogue comprises a dataset of some 230 digital editions, currently, with some fifty consistent and comparable pieces of data on each, that range from the edition's subject matter and URL, to its features, textual encoding scheme, and technological infrastructure. While it takes a more object, data-focused approach to reviewing and cataloguing the included editions, the Catalogue also uniquely offers the possibility of rich comparison and analysis across publications, even if that more subjective and evaluative work is yet to be done. It may also someday provide a model to be applied to the evaluation of other types of digital publishing projects, specifically like the Mellon-funded university press projects, the Getty's OSCI collaborative, and other open access, scholarly editions which have been the subject of our history here thus far.

(Albers, 2017)

The analysis in the present paper does not consider the fourteen digital editions in the Catalogue marked as *digitised* editions (i.e. whose value in the 'Digital' column of the Catalogue table is

---

<sup>16</sup> Available at: <https://web.archive.org/web/20170907125945/http://www.digitale-edition.de>

<sup>17</sup> See: [https://github.com/gfranzini/digEds\\_cat](https://github.com/gfranzini/digEds_cat)

<sup>18</sup> See: [http://rzblx10.uni-regensburg.de/dbinfo/detail.php?bib\\_id=allefreien&colors=&ocolors=&lett=k&tid=0&titel\\_id=10204](http://rzblx10.uni-regensburg.de/dbinfo/detail.php?bib_id=allefreien&colors=&ocolors=&lett=k&tid=0&titel_id=10204). The number of libraries that make use of this service increases on a monthly basis.

<sup>19</sup> See: [https://github.com/gfranzini/digEds\\_cat/issues?q=is%3Aopen+is%3Aissue+label%3Afeature](https://github.com/gfranzini/digEds_cat/issues?q=is%3Aopen+is%3Aissue+label%3Afeature)

0) or, in other words, projects that exist as electronic or digital versions of print editions.<sup>20</sup> This means that only 242 of the 256 editions present in the Catalogue were analysed.<sup>21</sup>

## Methodology

To answer this study's research question (*what are the expectations of digital editions in the Digital Humanities and how do these correlate with existing digital editions?*), on 30th March 2017 the authors circulated a web survey entitled 'Expectations of Digital (Textual) Editions' to collect information about what *users* expect or want from a digital edition. This survey was targeted at the Digital Humanities community,<sup>22</sup> the field the authors identify with and within which many digital editions take shape.

The decision to collect data about the users of digital editions through a *web* survey was dictated by a number of factors. Firstly, web surveys are inexpensive; secondly, they are often quicker to set-up and run compared to other methods, which typically require calls for participation and participant selection; and thirdly, they allow to survey a large population and to collect large data samples (Horner, 2011, p. 956). On the other hand, the nature of web surveys automatically excludes participants who do not have access to an Internet connection (*ibid.*). This drawback, however, was not relevant for this particular survey, as an Internet connection is the prerequisite for the consultation and use of digital editions,<sup>23</sup> whether users are accessing the web from their own machines or from library workstations. Those who do not have access to an Internet connection are unable to work with digital editions (unless these allow for offline use, which is not usually the case).

## Survey Design

In order to compare the results of the survey against the 242 digital editions in the Catalogue of Digital Editions, the questions were modelled upon the 49 cataloguing features used in the Catalogue. To encourage as many responses as possible, the authors opted for twenty questions only, some of which grouped features by typology or purpose (e.g. Question 11 brings together indices, string search, advanced search and APIs - see further below). The questionnaire

---

<sup>20</sup> The reader may ask why a digital edition is in the Catalogue if it is neither digital nor an edition (if, in other words, its 'Digital' and 'Edition' values in the Catalogue are set to 0). This is because, while some projects self-define as *digital* and *editions*, the meaning of both terms does not match that of Sahle's definitions, which the Catalogue uses as a reference to catalogue entries (2008).

<sup>21</sup> The reader is reminded that this analysis was done in May 2017 and thus ignores digital editions added to the Catalogue from June 2017 onward.

<sup>22</sup> The survey was distributed via the authors' Twitter accounts, and to the following mailing lists and Facebook groups: *Associazione per l'Informatica Umanistica e la Cultura Digitale (AIUCD)* at <https://web.archive.org/web/20170907130136/http://lists.lists.digitalhumanities.org/mailman/listinfo/aiucd-l>; *Corpora* at <https://web.archive.org/web/20170907130215/http://clu.uni.no/icame/corpora/sub.html>; *Digital Humanities im deutschsprachigen raum (DHD)* at <https://web.archive.org/web/20170907130307/http://dig-hum.de/dhd-mailingliste>; *Digital Classicist* at <https://web.archive.org/web/20170907130356/https://www.jiscmail.ac.uk/cgi-bin/webadmin?A0=DIGITALCLASSICIST>; *Digital Medievalist* at <https://web.archive.org/web/20170907130454/http://listserv.uleth.ca/mailman/listinfo/dm-l>; *Humanist* at <https://web.archive.org/web/20170907130536/http://dhhumanist.org>; TEI-L at <https://listserv.brown.edu/archives/cgi-bin/wa?A0=TEI-L>; *AIUCD* Facebook group at <https://www.facebook.com/groups/aiucd/>; *Digital Medievalist* Facebook group at <https://www.facebook.com/groups/49320313760/>; *European Association for Digital Humanities (EADH)* Facebook group at <https://www.facebook.com/groups/109971049335068/>; *UCL Centre for Digital Humanities* Facebook group at <https://www.facebook.com/UCLDH/>. Facebook and Twitter posts were re-advertised by individuals in numerous other social spaces.

<sup>23</sup> Early "offline" examples of digital editions, such as the *Thesaurus Linguae Graecae* (TLG), were published on CDROM, but these are rare (and often replaced, like the TLG, by an online version), difficult to access or expensive to purchase.



contained a mix of multiple choice and Likert scale questions<sup>24</sup> for a total of twenty questions. The Likert Scale was used to measure the *importance* of a particular feature in a five-point scale:

- 1 - Not important
- 2 - Slightly Important
- 3 - Moderately Important
- 4 - Important
- 5 - Very Important

After generally imposing a structure, Question 20 (*Is there anything else you would like to tell us about your user needs for digital editions that we have not covered here?*) was designed to give respondents the opportunity to freely express their views, thus mitigating what Krippendorff identifies as a common problem in such surveys:

For efficiency's sake, researchers gain a considerable advantage if they can impose a structure on the data-making process so that the results are readily analyzable. Surveys, mail questionnaires, and structured interviews typically offer respondents predefined choices that are easily tabulated, coded, or processed by computer. But they thereby also prevent the respondents' individual voices from being heard. (Krippendorff, 2004, p. 41)

The software chosen for this particular web survey was *Opinio*<sup>25</sup>, which is freely provided by the author's home institution. The survey broadly defined a digital edition as:

an edition or transcription of a text that makes use of digital technologies to enhance users' access and experience of the source material. It can be a completed or an ongoing project, and does not necessarily have to provide a critical commentary. A mere replica of a printed edition is, in this instance, not considered to be a digital edition.

As Horner reports (2011, p. 957), "[...] Web surveys may pose additional participant confidentiality issues", which might deter participants from contributing their answers. In this survey, demographic information about the participants, such as gender, age, ethnicity, location and religion, was not collected, as it was not considered relevant. The complete anonymity was announced in the survey's advertising message and the chosen software was set to not reveal any information about participants to the authors.

On the closing date one month later, 30th April 2017, the survey recorded 218 *completed* responses and 130 *incomplete* responses, for a total of 348 *stored* responses. Comparatively, the LAIRAH study recorded 149 completed responses over a period of four months (Warwick et al., 2008, p. 385) and Porter's aforementioned 2002 and 2011 surveys recorded 43 and 169 completed responses respectively.

With every survey, there is the issue of representativeness, that is, the quality or relevance of the results and if these adequately represent the total target population (Schouten et al., 2009). Despite being set-up to *qualitatively* assess digital editions in relation to user expectations, the survey's large number of responses might also be considered statistically relevant. However, as no evidence exists with regard to population size, the authors refrain from claiming

---

<sup>24</sup> For more information about the Likert scale, see Allen and Seaman (2007).

<sup>25</sup> Available at:

<https://web.archive.org/web/20170907130835/https://www.ucl.ac.uk/isd/services/learning-teaching/elearning-staff/core-tools/opinio>. *Opinio* provides survey data in HTML, PDF, SPSS and Raw Data formats.

representativeness, and acknowledge the qualitative nature of the results. While preliminary, these results are still valuable for those who are planning to build digital editions.

In the next section, we juxtapose the features of 242 digital editions built by the community present in the Catalogue of Digital Editions with the 218 *complete* answers collected through this web survey in order to identify meeting and diverging points. The 130 incomplete questionnaires are not considered in this discussion, as their differing degree of completeness do not make for a uniform analysis. Median (of the five-point Likert Scale) and Standard Deviation (from the median) values will be hereafter referred to as *M* and *SD* respectively.

## Results

To facilitate the presentation of the statistics, the present article does not defer the discussion of the results but divides them between the twenty survey questions to readily contextualise them. Therefore, where applicable, for every question we provide a breakdown of the corresponding responses, their correlation with the statistics gathered from the Catalogue of Digital Editions, a brief discussion and ensuing practical recommendations. The results section is then followed by a summary discussion, which visualises the overall creation trend in digital editing against the results of this survey.

### Questions 1, 2, 3: Overview of Respondents

The largest group within the 218 survey participants identified themselves as a researcher (38.67%); professors follow in second place (25.78%); the undefined 'other academic position' participants place third (12.5%); and students in fourth (11.72%); and 29 participants provided new identifying categories, some of which could have been represented by the options already provided (e.g. the free-text answer 'PhD student' was not necessary seeing as a 'student' category was already provided). This is a first indication that digital editions are more likely to be accessed and potentially reused in research.

Almost 82% of participants identified themselves as belonging to the Humanities, 10% as working across different disciplines and 7.85% as belonging to the Applied Sciences. These results may seem unsurprising, given that the documentary and literary content of digital editions is at the core of Humanities studies, but we have no concrete information as to what exactly forms this clear majority. There could be many reasons for the survey's popularity among humanists: the advertising channels used to circulate the survey were mostly humanities-oriented and would therefore explain the low science participation; traditionally, the preparation of (digital) editions of documentary heritage is a humanities-driven activity; the low presence of scientists might also be dictated by the fact that few practitioners have a need for digital editions of cultural and historical texts.

With respect to the specific disciplines of the 218 responding participants, 44% are involved in literary studies and 22% conduct historical studies. Within this 66% subpopulation of literary and historical studies, 17% focus on the classical period, while 16% research upon the Middle Ages. Other disciplines are as diverse as Shakespeare Studies, Law, History of Ideas, History of Daily Life, Television, Agriculture, Medieval Legal History, Medieval Medicine, Surgery and the History of Medicine, Musicology, Liturgical Medieval Chant, History of exotic animals in the Middle Ages, Folklore, Engineering, Historical Metrology, Theatre, Architectural History and Anthropology, Military History, Hittitology, Children's Literature, Language Pedagogy, Food History, Digital Archaeology, Sinology, Lexicology, Social History of the Ottoman Middle East, Conceptual History, Journalism, Visual Rhetoric, Caribbean Literature, Celtic Studies, Information Science, Sociology of culture and Science, History of Photography and Bureaucracy. Those who identified as scientists include, among others, a computer scientist focussed on interaction design, a cyber security and automation engineer, an information

architecture writer, a data modeller and web frontend developer, and a computer programmer with an interest in classical literature.

#### Question 4: What do you primarily seek in a digital edition?

Participants were asked to specify what primary use they make of digital editions. The largest group of users, 79 (36%), consume digital editions for data and *public* reuse, while 70 (32%) users look for a complete educational resource to learn more about the subject matter. A third group of users, 55 (25%), is interested in digital editions for *private* (re)use. One user stated that (s)he seeks the "reconstruction of the lost original of a text" and all materials used to that end; another user expressed the need to annotate digital editions and share the annotations; two users emphasised the importance of being able to search a text.

#### Question 5: *How important is the scholarly component of a digital edition for you?*

This question was designed to elicit the rate of importance of the *scholarly* component of a digital edition. According to Sahle's definition of a digital scholarly edition, used in this survey as a reference for answers, projects that do not offer a critical examination of the text(s) are not scholarly (2008).

**Survey results (M = 4, SD = 1.02).** Almost 50% of respondents rated the scholarly component of a digital edition as *very important*. Despite being affected by the three 'No opinion' outliers, the Median score indicates that there is an expectation that digital editions should critically assess the texts at hand.

**Catalogue data.** Of the 242 digital editions under consideration in the Catalogue, 29 (11%) were catalogued as not scholarly.

**Recommendation.** This low percentage (11%) is encouraging as it suggests that creators of digital editions are already meeting the needs of many users in this regard. Creators are reminded to always provide a glossary or list of domain-specific terminology in their critical edition, regardless of its intended audience.

#### Question 6: *How important is knowing the production lifespan or duration of a digital edition (begin and end years)?*

This question was asked to understand how important it is for users to know the production span of a digital edition, that is, the time taken to *complete* the digital edition as intended by its editors. This duration can vary greatly: one ongoing project will be funded for as long as fifteen years<sup>26</sup>, while a completed project ran for sixteen years.<sup>27</sup> Duration is not necessarily indicative of the quality of a digital edition but combined with other factors, such as funding and team size, it provides a rough understanding of what is achievable with a given amount of resources, therefore offering a means of comparison. Start and end dates also place digital editions in a historical and chronological context, helping users better appreciate the technological affordances and choices made at a particular time.

**Survey results (M = 4, SD = 1.23).** 63 (28.9%) participants rated the knowledge of the duration of a digital edition as *important*, while slightly more than 23% of participants consider it to be a *very important* characteristic.

**Catalogue data.** Of the 242 digital editions under consideration in the Catalogue, 69 (28.5%) do not provide either begin or end years ('forthcoming' is included in this count as it does not

---

<sup>26</sup> The *Edition Humboldt Digital* project, for example, expects to reach completion in 2032. See: <https://web.archive.org/web/20170907130923/https://dig-ed-cat.acdh.oeaw.ac.at/editions/detail/251>

<sup>27</sup> The Dead Sea Scrolls project, at: <https://web.archive.org/web/20171008134728/http://dss.collections.imj.org.il>

provide a more specific estimate of the expected project start date); 120 (49.5%) provide both begin and end years (where the end year can also be 'present', e.g. 2017); and 53 (21.9%) provide partial information by either specifying the beginning *or* the end year of the project. In sum, just over half of the projects in the Catalogue (50.5%) provide complete information about project duration.

**Recommendation.** Accordingly, given this question's high Median of 4, it is recommended that creators publish beginning and (expected) end years.

### Question 7: *How important is knowing which audience the digital edition targets (e.g. students, scholars, general public)?*

Although creators do not preclude unintended audiences from using their resources, making this information known helps users contextualise the data and better to understand the objectives of the creators. If, for instance, a digital edition is targeted at *experts* of medieval Latin manuscripts, creators might perceive the publication of a glossary of terms or a list of conventions as redundant, and thus refrain from making one available. Making this choice abundantly clear on the project website helps define the scope of the endeavour, and guides users accordingly.

**Survey results (M = 3, SD = 1.19).** Almost 31% of respondents rated target audience as *moderately important* information, 23% as *important* and 22% as *very important*.

**Catalogue data.** Of the 242 digital editions considered in the Catalogue, the majority, 141 (58.2%) do not provide information about the intended audience. Out of the remaining 101, 54 (53.4%) explicitly target the general public (analogous terms used include 'global audience' and 'laypeople').

**Recommendation.** Given the response to this question, it is recommended that creators specify the intended audience of their project.

### Question 8: *How important is detailed editorial and technical documentation (i.e. glossaries, information about the technologies used to produce the digital edition, the imaging settings, was the text OCR'd or keyed, etc.)?*

Documenting the process of creation behind any type of project serves to communicate development, quality and to give appropriate context. With documentation, creators can make aims, limitations and the expected use known to their users, as well as facilitate the reuse of a resource. One of the basic principles and assumptions of research is *reproducibility* or, in other words, the ability of one researcher to take the work of another researcher, follow their pathway and arrive at the same results. Reproducibility is key to research acceptance and validation, so much so that entire courses are based on this principle.<sup>28</sup> While reproducibility might not be essential to all disciplines (Casadevall and Fang, 2010), scholars argue that research should be sufficiently documented in order for it to be accepted by the community.<sup>29</sup> Moreover, although scholars in the Humanities do not typically reproduce digital editions, some of their constituent parts might be reusable (e.g. the texts themselves, an XML schema or other underlying code), and should, therefore, be adequately described to facilitate reuse and reproducibility (Allison, 2016). Ten years ago, the LAIRAH study found that many digital humanities resources did not keep organised documentation (2008, p. 391).

---

<sup>28</sup> Such as the Coursera *Reproducible Research* online course by Johns Hopkins University, available at: <https://www.coursera.org/learn/reproducible-research>.

<sup>29</sup> A 2016 *Nature* survey asked 1,500 researchers to measure reproducibility, with somewhat alarming results. See: <https://web.archive.org/web/20170907131248/http://www.nature.com/news/1-500-scientists-lift-the-lid-on-reproducibility-1.19970>.

**Survey results (M = 4, SD = 1.12).** An aggregate of 152 participants (69%) rated the importance of detailed editorial and technical documentation between *important* (24.31%) and *very important* (45.41%).

**Catalogue data.** Of the 242 digital editions considered in the Catalogue, four of them do not make clear what the source of the edited text is (i.e. is it derived from a printed edition, what is the base text, on which digitised documents is the digital edition based on, is it a new born-digital edition), while 238 provide some form of philological or editorial statement. Of the 238 projects, 122 (51%) provide partial information with regard to the source of the text and the editorial policy, and the remaining 116 (49%) provide complete information.

With respect to the technical information, that is, information about technologies used and related standards used, 118 (48%) digital editions do not provide any such information, 39 (16%) provide partial information and 68 (28%) provide complete information; finally, the websites of the 17 (7%) digital editions published on CD-ROM or protected by a pay-wall/login system also do not provide this information, which is not to say that it is not available on the CD-ROM itself or within the subscription portal.

Only 50 projects (approximately 20%) out of 242 provide both complete editorial and technical documentation. Disappointingly, these results corroborate the LAIRAH findings published over ten years ago.

**Recommendation.** Given the low statistics currently discernible in the Catalogue and, conversely, the high rating of this question, creators must take more care to incorporate comprehensive documentation. This should provide descriptive information about the project itself, including purpose, motivation, duration, limitations, human and financial resources invested, and target audience. Additionally, and also in reference to Question 9 below, it should provide details about the source document's history, its significance, context, its provenance and current repository or location, as well as technical documentation pertaining to its digitisation (image-capturing equipment and settings) and to the entire project back-end.

*Question 9: How important is the exhaustiveness of contextual information (e.g. current repository of source material, links to external resources, quality of source materials, etc.)?*

Similarly, participants were asked to rate the importance of contextual information, such as the provenance and current repository of the source material(s).

**Survey results (M = 4, SD = 0.93).** An aggregate of 160 participants (73%) rated the provision of contextual information between *important* (36.24%) and *very important* (37.16%).

**Catalogue data.** Of the 242 digital editions considered in the Catalogue, 191 (78.9%) provide information about the institutions currently housing the source material(s), while 19 (7.8%) do not and 32 (13%) are built on materials that do not have a specific physical location (e.g. print editions published in multiple copies). With respect to provenance, only 6 projects (2%) do not provide clear information, one project does not apply and the remaining 235 (97%) either provide the countries or the cities of provenance. Finally, 121 (50%) provide links to external resources and supplementary materials, and 121 (50%) do not.

**Recommendation.** For this question, the Catalogue shows that, generally, projects are providing contextual information and are thus meeting the expectations expressed by the participants.

*Question 10: How important is the provision of high quality digital images upon which the digital edition builds?*

**Survey results (M = 4, SD = 1.12).** This question found a high consensus, with an aggregated 140 (57.8%) participants rating the provision of digital images to accompany the edited or transcribed texts either *important* (26.15%) or *very important* (38.07%).

**Catalogue data.** Of the 242 projects under consideration, 135 (55.7%) provide images, 85 (33.8%) do not and 13 (5.3%) provide only some. The remaining 12 are catalogued as 'not provided', a value used for projects published on CD-ROM or behind a pay-wall (inaccessible to the authors) to indicate that the information available on the site is inconclusive with regard to the provision of images. Of the 135 projects (55.7%) that provide images, 107 (79%) allow users to zoom in and out, and 13 (9%) come with text-image linking functionality to enhance the reading experience of the document.

**Recommendation.** The response rate to this question suggests that many users expect digital editions to provide images. The provision of images, however, is often tied to copyright regulations enforced by rights holders (e.g. individuals and/or institutions), and these do not always facilitate access or employ fair use of their resources in teaching, learning and scholarship.

There is the demon of copyright. Some of the most exciting digital edition projects focussed on modern authors. It can be difficult enough gaining permission for print editions for these; for digital editions, in some notorious cases, it has proved impossible. But even for older texts, where there should be no copyright issues, there have been problems. Arranging for digital photography and reproduction rights is, with very rare exceptions, arduous and too often forbiddingly expensive.

(Robinson, 2010)

In light of these restrictions, creators of digital editions must endeavour to establish and clear rights on the images they plan to use or secure publication permissions *before* the start of the project. If permission is not granted by the image holders and the digital edition proceeds without images, creators are strongly advised to publish a visible statement on the project website documenting this drawback. The negative publicity to the rights holders will not only increase the awareness of the issue but will also clearly communicate to users the imposed limitations of the project.

### Question 11: *How important is the availability of advanced functionality and browsing, such as indices, filtering, searching and Application Programming Interfaces (APIs)?*

**Survey results (M = 4, SD = 0.92).** The great majority of participants (84%) rated the browse and search functionalities of digital editions, such as indices, text searches and advanced filters either *important* (34.4%) or *very important* (50%). Application Programming Interfaces (APIs) were included in this category as they provide a similar means of understanding how the components of a digital edition come together "under the hood". Four participants (1.8%) rated search functionalities as *not at all important* and the remaining 14% rated them as *slightly* to *moderately important*.

**Catalogue data.** Of the 242 digital editions considered in the Catalogue, 12 (4.9%) come with an API, 109 (45%) provide indices, 152 (62%) provide a text or string matching search and 113 (46%) provide advanced search functionality. A total of 100 (41%) projects provide both string search and advanced functionality.

**Recommendation.** Given this question's high Median of 4, it is recommended that, where applicable, creators provide search functionality (string searches, indices, filters, concordances and more advanced options) to help users locate information more easily and, in the case of APIs, repurpose and integrate information in other resources.

### Question 12: *How important is the possibility of consulting a digital edition's website in multiple languages other than English?*

Digital editions are often the result of international collaborations and in an effort to maximise outreach, some projects localise the website and interface of the digital edition.

This question sought to elicit the importance of providing multilingual sites, *not* translations of the content of digital editions.

**Survey results (M = 2, SD = 1.33).** The high SD value of this question indicates that there is no clear consensus, with 25% of participants rating multilingual websites or web interfaces as *not important* against an aggregate 28% who judge it between *important* (17.89%) and *very important* (10.09%).

These results attest to English as the recognised global *lingua franca* (Montgomery, 2013; Seidlhofer, 2011), and yet a sizable portion of participants expressed the need for a language option. Given that the survey did not collect any personal information about participants, it is impossible to prove or disprove hypotheses about this result based on the nationality of the respondents. Even so, the rating for this question might not necessarily be dictated by the respondents' own linguistic abilities (i.e. an English-speaker might not consider multilingual sites important as many are published in English anyway) but could also constitute an independent evaluation of what they think ought to be done in this regard. Responses may have also been influenced by the modern availability of tools to translate web pages (e.g. Google Translate's browser extension for Google Chrome<sup>30</sup>); despite their limitations as machine-translators, these tools can, at the very least, help non-natives locate information more easily.

**Catalogue data.** Looking at the Catalogue, of the 242 digital editions considered, 210 (86.7%) project websites are published in one language only and, of these, 158 (75.2%) are in English. The remaining 32 (13.2%) projects are published in two or more languages and all 32 provide an English translation or domain.

**Recommendation.** In light of the mixed opinions recorded in this survey with regard to multilingual sites, and unless this feature is explicitly required by funding agencies, creators might want to reconsider providing localised versions of their resources (i.e. translations of the user-interface of the edition) or prioritise these differently in project development. While localisation opens up resources to a wider audience, the outcome of this survey suggests that the resources needed for translations and their maintenance might be better invested in features rated as very important and thus more useful to users.

### Question 13: *How important is the provision of data in Open Source/Access formats?*

The Digital Humanities has joined the world movement to make scientific research available to the widest possible audience in free and open form (Hamilton and Saunderson, 2017). Community experts have addressed the importance of open source and open access models in Digital Humanities practice,<sup>31</sup> and this survey question sought to hear the opinions of users in this regard. The Catalogue of Digital Editions categorises openness into five levels:

- Proprietary, all material is copyrighted. The source is closed and not reusable by other research projects. To access the material, users must pay a subscription fee.
- Same as above but the subscription is free of charge.
- Open Access. The texts may be accessed through specific software but the source is not accessible.

---

<sup>30</sup> See: <https://chrome.google.com/webstore/detail/google-translate/aapbdbdomjkkjkaonfhkkikfgjllcleb?hl=en-GB>

<sup>31</sup> See, for example, Cohen, D. (2010), Ramsay, S. (2010) and Fitzpatrick, K. (2010).

- Open Access and Partial Open Source. Part of the data underlying the digital edition (e.g. text but not images) is freely available for access and reuse.
- Open Access and Open Source. All data underlying the digital edition is freely available for access and reuse.

**Survey results (M = 5, SD = 0.93).** The vast majority of participants (77%) rated the importance of Open Source and Open Access formats between *important* (22%) and *very important* (55%). Only 3 participants rated it as *not at all important* and 7 participants expressed no opinion.

**Catalogue data.** Of the 242 digital editions considered in the Catalogue, one project has yet to make data available, 32 (12%) are protected by a pay-wall, 9 (3.7%) are accessible through a free registration process, 125 (51%) allow data to be accessed but the source is not accessible, 49 (20%) are both Open Access and Open Source but only part of the source is available for download and reuse; finally, 26 (10%) are both Open Access and Open Source making all of the source available for download and reuse. Furthermore, only 63 (26%) projects release their data under various forms of Creative Commons Licenses.

[...] The major problem I see, in practice, is the strange reluctance lots of TEI projects still have to expose their TEI source directly.

(Lou Burnard, 23 November 2017)<sup>32</sup>

**Recommendation.** Given that this question has the highest possible Median of 5, it is strongly recommended that creators adhere to Open Source/Access policies to the fullest extent possible, and that permissions be made clear on the project website (i.e. which parts of the project can be reused and under what conditions). While Creative Commons Licenses are typically used for non-software works, there are many other permissive licences to choose from depending on the needs of the project, including the Academic Free Licence (AFL), MIT, Apache and GPL Licences.<sup>33</sup>

**Question 14:** *How important is knowing the financial and human resources invested in the production of a digital edition (e.g. amount of funding obtained, number of researchers/staff involved)?*

Large teams are not necessarily indicative of quality as they may not cover the range of skills required to produce a digital edition (Warwick et al., 2008, p. 387).

**Survey results (M = 3, SD = 1.17).** Although human and financial resources are not necessarily indicative of the quality, or lack thereof, of a digital edition, providing some information in this regard is considered by almost 31% of participants to be *moderately important*.

**Catalogue data.** Of the 242 digital editions considered in the Catalogue, 223 (92.1%) do not provide information concerning the budget or grant size, 12 do, and 7 do not apply as these projects are carried out as leisure activities. The Catalogue does not currently record information pertaining to the size of the teams behind digital editions, so no numbers are available at this time.

**Recommendation.** It is recommended that creators of digital editions include information about team and grant size in the project description as this helps users gain a better understanding of the scope of the project and of how resources influence its development.

---

<sup>32</sup> Excerpt taken from a discussion held on the *Text Encoding Initiative* (TEI) mailing list:  
<https://listserv.brown.edu/archives/cgi-bin/wa?A2=TEI-L;3604f20d.1711>

<sup>33</sup> A comprehensive list of open source licences can be accessed at:  
<https://web.archive.org/web/20170907132405/https://opensource.org/licenses>



**Question 15: *How important is the mobile-device compatibility of a digital edition (i.e. the possibility of using it on a tablet and/or smartphone)?***

Web technologies are increasingly adapting desktop web browsing to a variety of mobile devices in order to make content as widely accessible as possible.<sup>34</sup> This survey question sought to understand users' views on accessing digital editions from handheld devices.<sup>35</sup> The survey did not ask whether users make the *fullest* use of digital editions on their handheld devices, preferring instead to focus on *access* intended as light browsing.

**Survey results (M = 3, SD = 1.34).** As indicated by the high SD, respondents took a range of views: 50 users (22.9%) rated mobile-compatibility as *slightly important* and another 50 (22.9%) as *important*; 36 (16.5%) users rated it as *not at all important* and 35 (16%) as *very important*. Of the remaining 47 users, 3 expressed no opinion and 44 (20.18%) rated it as *moderately important*.

**Catalogue data.** The mobile compatibility of digital editions present in the Catalogue is verified using Google's *Mobile-Friendly Test*<sup>36</sup>. Of the 242 projects considered in the Catalogue, 39 projects (16%) passed the Google mobile-friendly test while 203 (83%) did not.

**Recommendation.** While the responses recorded by this web-survey with regard to the importance of mobile access to digital editions was not conclusive, employing technologies to this effect is recommended if creators intend for their project to be more widely visible and usable. More importantly, the adherence to web accessibility standards, such as those issued by the World Wide Web Consortium<sup>37</sup>, to champion inclusivity should become an integral part of the development process of a digital edition. The findings of a survey entitled "Inclusive Design and Dissemination in Digital Scholarly Editions", circulated in July 2017, will tell us more about the types of access and web accessibility options of digital editions observed by the scholarly community.<sup>38</sup>

**Question 16: *How important is the possibility of downloading and reusing the data published within a digital edition?***

**Survey results (M = 5, SD = 0.85).** Consonant with Question 13 about Open Source/Access, the majority of participants (57%) rated the possibility of downloading and reusing digital edition content as *very important*. One participant did not express an opinion and none of the participants rated this feature as *not at all important*.

**Catalogue data.** In its present form, the Catalogue only records the availability of downloads where the digital edition is encoded in XML(-TEI). Of the 242 digital editions considered, 132 (54.5%) are encoded in XML(-TEI) and, of these, 48 (36%) allow users to download the XML(-TEI) files. Only a handful of projects –the Catalogue does not specify the exact number yet but this feature will be added in the near future– provides a single bulk download option for the entire digital edition. As expressed by one participant in answer to Question 20 (see further below), bulk downloads are desirable and a more efficient means of reusing data.

**Recommendation.** These statistics show that the digital editions examined are not adequately meeting the download expectations expressed by the participants of this survey. It is therefore strongly recommended that creators enable download and preferably in bulk.

---

<sup>34</sup> For instance, the mobile-responsive framework *Bootstrap*, available at <https://web.archive.org/web/20170907132507/http://getbootstrap.com/>, or the *WordPress* Content Management System, available at:

<https://web.archive.org/web/20170907042719/https://wordpress.org/>

<sup>35</sup> For a study on the dissemination of digital scholarly editions via mobile devices, see (Kelly, 2015).

<sup>36</sup> Available at: <https://web.archive.org/web/20170907132703/https://search.google.com/test/mobile-friendly>

<sup>37</sup> See: [https://web.archive.org/save/\\_embed/https://www.w3.org/](https://web.archive.org/save/_embed/https://www.w3.org/)

<sup>38</sup> Available at:

<https://web.archive.org/web/20170907132815/https://www.surveymonkey.com/r/MCDRMYY>

**Question 17: *What use would you make of the data published in a digital edition?***

This question sought to learn more about the types of (re)use some users make of data published in digital editions. *Teaching* was placed first with a frequency of 31%, closely followed by *text analysis* with a rate of 30% and *corpus aggregation* or building with 21%. The free-text input recorded other uses, including *research*, *literary analysis*, *re-editing* and *annotation*. This variety of applications speaks to the value of digital editions not only as research-enabling instruments but also as pedagogical tools worthy of being used in the classroom alongside more traditional study materials.

**Question 18: *Which of the following data formats would facilitate your studies?***

Many digital editions adhere to XML(-TEI) standards to encode texts. Editors especially advocate this practice as a suitable means of marking-up and publishing texts online. Despite TEI's claims to interoperability, the fact that the selection and use of an XML(-TEI) tag is based on the human interpretation of that tag inherently obstructs an effortless reuse of XML(-TEI) files (Schmidt, 2014). It follows that file reuse requires some form of adjustment to fit the new context or purpose. In some cases, adjustments can turn into extensive pre-processing tasks in order to get the data into a (re)usable format.<sup>39</sup> To incentivise, and reduce the complexity of, reuse, some creators of digital editions, and of digital textual resources in general, present texts in multiple formats.

**Survey results (M = 4, SD = 2.12).** This multiple-choice survey question was designed to elicit participants' views on data formats in digital editions by asking them to state their preference. XML(-TEI) tops the list with a frequency of 19.31%, followed by images optimised for the web (e.g. PNG and JPG) (17.66%), Plain Text (16.14%), PDF (15.86%), XML (12.97%), TIFF images (11.59%) and ePub (4.69%). Free-text replies included Microsoft Word (3 participants), CSV (1 participant) and JSON (1 participant).

The results obtained for images contradict Robinson's claim:

Firstly: it appears that rather few readers (indeed, rather often, only the editors) actually want to see all the images, all the transcripts, all the collations. Traditional print editions acted as filters, straining out all this information so that readers did not have to see it: if readers do not want to see it, then including all this is no advantage at all.

(Robinson, 2010)

In this article, Robinson does not provide evidence for these claims but the results of this survey either contradict this article or suggest that the provision of high quality images over the past seven years has caused a change in user mentality and needs.

**Catalogue data.** Of the 242 digital editions considered in the Catalogue, 136 (56%) are encoded in accordance with XML(-TEI) standards, 86 (35%) do not use XML at all, and 135 (55%) texts are accompanied by images of the source documents.

The Catalogue does not list all available data formats for each project, so there are currently no numbers with regard to the projects that, for instance, also provide Plain Text (TXT) versions

---

<sup>39</sup> The TEI Wiki dedicates a page to 'Conversion and Preprocessing Tools'. See: [https://web.archive.org/web/20170907132852/https://wiki.tei-c.org/index.php/Category:Conversion\\_and\\_preprocessing\\_tools](https://web.archive.org/web/20170907132852/https://wiki.tei-c.org/index.php/Category:Conversion_and_preprocessing_tools)

of the edited texts. This information, planned as a future addition to the Catalogue, would prove particularly useful to those who wish to run different analyses on a text, as recently expressed in a *Digital Medievalist* mailing list thread:

[...] is there a way to access the plain text directly, or is there only a search interface at the moment? Having direct plain text access can be useful for others to do various further analysis on the corpus.

(Nick White, 29 June 2017)<sup>40</sup>

And, again, three years ago in the *Digital Humanities Questions & Answers* forum:

I'd like to use the available corpora in the German Text Archive (<http://www.deutschestextarchiv.de/download>) to train OCR software. For this I need these texts as plaintext. All the German Text Archive texts however are all TEI P5 tagged. How do I best convert these (hundreds..) of documents into plaintext?

I'm comfortable on the command line and with small shell scripts but I wouldn't be able to write an app to make use of a public API to such a service. Ideally I'd like to find some tei2text-ish command line tool but the ones I've found in googling around and looking on GitHub don't appear (to me, leastways) to be suitable for TEI texts.

(Arno Bosse, 2015)<sup>41</sup>

**Recommendation.** Based on the responses recorded for this question, it is recommended that creators provide a marked-up (XML) text for users interested, for example, in close-reading the text, as well as a Plain Text (TXT) version of the same text to meet the needs of those who wish to perform some form of computer-aided text analysis.

### Question 19: *Can you provide an example of a digital edition that has good functionality for your needs? Why and how does it meet your needs?*

The intention of this question was to draw out examples of digital editions that participants feel meet their needs (see Appendix A). The majority of participants provided examples of projects they consider satisfactory, specifying both the positive and their negative aspects:

Response 7. *I've found the search function of the Loeb Online very helpful and comprehensive, although obviously it's a pain in the bottom to navigate.*

Response 21. *Oldbaily [sic], criminocorpus...excellent information retrieval and statistical analysis. Download and reuse options could be better.*

Digital editions mentioned by multiple respondents are the *Perseus Digital Library* (seven respondents),<sup>42</sup> *Electronic Beowulf* (five respondents),<sup>43</sup> *Folger Digital Texts* (four

---

<sup>40</sup> See: <https://web.archive.org/web/20170906104222/http://listserv.uleth.ca/pipermail/dm-l/2017-June/003029.html>

<sup>41</sup> This question and all given answers can be viewed here:

<https://web.archive.org/web/20180221090725/http://digitalhumanities.org/answers/topic/how-do-i-best-convert-hundreds-of-tei-p5-documents-to-plaintext>

<sup>42</sup> Accessible at: <https://web.archive.org/web/20170907132947/http://www.perseus.tufts.edu/hopper/>

<sup>43</sup> Accessible at: <http://ebeowulf.uky.edu>; the project's record in the Catalogue is available at: <https://web.archive.org/web/20170907133128/https://dig-ed-cat.acdh.oeaw.ac.at/editions/detail/145>

respondents),<sup>44</sup> *Online Froissart* (three respondents),<sup>45</sup> *Loeb Classical Library* (two respondents),<sup>46</sup> *Corpus Corporum* (two respondents),<sup>47</sup> *e-codices* (two respondents)<sup>48</sup> and the *Bayeux Tapestry* (two respondents).<sup>49</sup>

Some participants, 22 in total, stated that no project currently meets their needs:

Response 30. *There is none.*

Response 42. *None are really good. Google Books is almost always bad.*

Response 78. *I'm not sure I have found one that meets my needs for teaching [...].*

Response 96. *I've never had the pleasure to use an edition that fulfilled my ideal.*

Six participants (Responses 1, 28, 38, 71, 106, 111) self-publicised their own work (but only responses 1 and 71 provided a link to the resource) and there is no way to tell from this survey whether other participants did the same in less overt ways. The claims of these six participants reinforce the aforementioned LAIRAH findings: expressing appreciation over one's own digital edition suggests that the digital edition was built to primarily fulfil the needs and requirements of the creator. This line of enquiry deserves further attention to determine the extent to which creators of digital editions engage with their target users during the preparation and development stages of the project.

The general impression one can glean from reading all of the answers given to this question is users' understanding that digital editions are imperfect tools and that they cannot meet the needs of every single user. Although creators may feel reassured by this awareness, the answers given in this survey carry a somewhat negative tone, suggesting tolerance towards the issue rather than acceptance.

### Question 20: *Is there anything else you would like to tell us about your user needs for digital editions that we have not covered here?*

With this question, the authors sought to fill any gaps in the survey by asking participants to express concerns or make suggestions with respect to their expectations of digital editions. The following were chosen as representative of the full answer-set (see Appendix B):

Response 3. *Much more important than fancy browsing, searching capabilities on the digital edition's site is the availability of either an API or the full XML-TEI download option [...].*

Response 10. *How can we have a digital edition tailored to various needs in one place?*

Response 13. *The best editions are about providing textual data to researchers, not dictating how researchers will read or make use of the data.*

Response 14. *The ability to download all XML files in an edition - preferably all in one single download - would be very useful. Not available in that many editions.*

---

<sup>44</sup> Accessible at: <http://www.folgerdigitaltexts.org>

<sup>45</sup> Accessible at:

<https://web.archive.org/web/20170907133344/https://www.hrionline.ac.uk/onlinefroissart/>; the project's record in the Catalogue is available at:

<https://web.archive.org/web/20170907133427/https://dig-ed-cat.acdh.oeaw.ac.at/editions/detail/124>

<sup>46</sup> Accessible at: <https://web.archive.org/web/20170907133457/https://www.loebclassics.com>

<sup>47</sup> Accessible at: <http://mlat.uzh.ch/MLS/index.php?lang=0>

<sup>48</sup> Accessible at: [https://web.archive.org/save/\\_embed/http://www.e-codices.unifr.ch/en](https://web.archive.org/save/_embed/http://www.e-codices.unifr.ch/en)

<sup>49</sup> Accessible at: <http://www.sd-editions.com/bayeux/zoom/#/facsimile%3DBayeux%26panel%3D1>

Response 22. *All the and only the essential information (name of the edition, authors, publishing institution, publishing date, last update, topic/mission of the edition) should be visible on the landing page. No other distraction.*

Response 32. *Digital editions are not just of interest to and production [sic] by those in literature and cognate areas!*

Response 38. *Clear and open license, provide a manual so theat [sic] users can actually RTFM.<sup>50</sup>*

Response 45. *Stable (persistent) URLs for resources; clear version system (=defined versions/updates of the digital edition).*

Response 57. *You mention the work done to put the edition together: no-one I should think thinks the technical and editorial labour is worth recording but it should be: only by recording it as standard will we get this work credited and the resources we need to produce it made available.*

Response 58. *Always, always, always evaluate how accessible your digital edition is for people with disabilities.*

Response 74. *A clear License Statement is the most important thing a digital edition must have. Other: Are the Responsible Contacts for the edition available/named?*

Response 75. *A standard for critical apparatuses which show more than the reading variants is desperately needed.*

## **Discussion and Recommendations**

How do the digital editions in the Catalogue compare to the responses recorded by this survey? The correlation is best captured as a chart, as shown in Figure 1: in this histogram, the x-axis lists the features of the digital editions that survey participants were asked to rate; for every feature, the histogram provides five coloured bars, each corresponding to a Likert scale point ('5-important' being 'Very important' and '1-important' 'Not important'). The overlaid red point-line plots the percentage of digital editions in the Catalogue that possess that particular feature. Generally, the bigger the distance between the red points and the tip of their respective coloured bars, the better the "performance" of the digital editions with respect to survey responses. For example, the highly rated scholarly component of digital editions is present in approximately 88% of the projects in the Catalogue, showing that digital editions are adequately meeting user needs. In contrast, digital editions in the Catalogue are performing badly when it comes to documentation, with only 50 projects (approximately 20%) providing both editorial and technical documentation, and to access, with only 26 projects giving users full access to the source. Moreover, despite receiving the highest rating in the survey, only 36% of digital editions in the Catalogue provide download options.

---

<sup>50</sup> Abbreviation of "Read The Fucking Manual", a disparaging comment used often on social media, see:  
<https://web.archive.org/web/20170907134134/http://www.urbandictionary.com/define.php?term=RTFM>

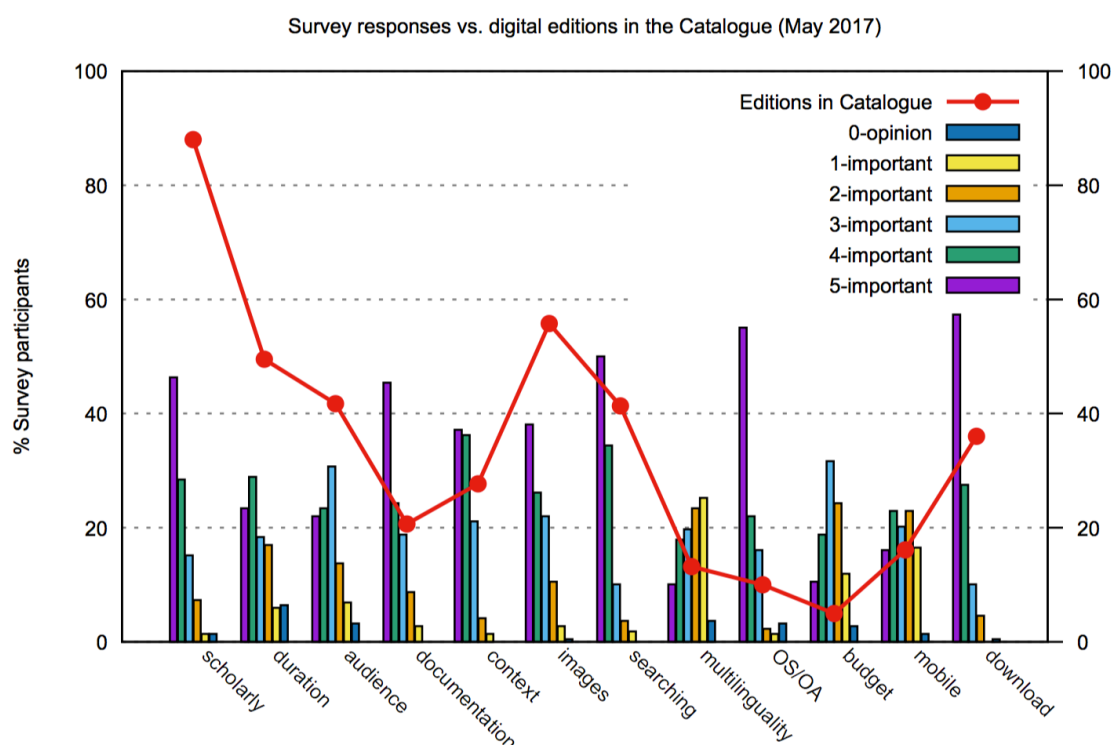


Figure 1. In this histogram, the x-axis lists the features of digital editions survey participants were asked to rate; for every feature, the histogram provides five bars, each corresponding to a Likert scale point ('5-important' being 'Very important' and '1-important' 'Not important'). The overlaid red point-line plots the percentage of digital editions in the Catalogue that possess that particular feature.

The conclusion one can draw from the results discussed is that the digital editions collected in the Catalogue of Digital Editions only adequately cover roughly half of the features examined in this study. Creators of these digital editions need not take these results as a denunciation of their efforts but, rather, as an invitation to reflect on how their editions can be improved to meet user requirements. To promote usefulness and fight the risk of neglect, funding agencies can support creators by formalising mandatory requirements and by allocating funds to the necessary administrative assistance that each project requires. To this end, the authors propose four deliverables –modelled against the LAIRAH recommendations (2008), the responses recorded in the present survey as well as the data collected in the Catalogue– that they believe should become standardised in grant application forms. These are:

**Staff training.** Shortcomings of digital editions can sometimes be traced back to the absence of a particular skill within the team. For this reason, funders should allocate sufficient resources for the training of Research Staff should no candidates with the optimal set of skills be available. Training should take place at the beginning but also during the project, depending on the employment status of the research team.

**User surveys and contact.** Users must have a voice in the development of a digital edition. To strengthen their role, funders are advised to make user studies a mandatory deliverable. These studies should iteratively evaluate the progress of the project against the needs of the users, so that, if possible, any modifications and additions can be factored in *during* the development of the project. Contact details should also be a compulsory requirement to offer users a means of communication with creators.

**Maintenance.** Dissemination and regular activity have been shown to reduce the neglect of a resource (Warwick et al., 2008, p. 389). These, however, can be time-consuming and are typically carried out by the researchers themselves, robbing the project of valuable research time. To help digital edition teams make the best use of research time, funders should allocate resources to hiring dedicated staff –on part-time positions even– to cover the marketing, management and administrative obligations of digital edition projects.

**Documentation.** Documentation is key to communicating the quality and value of the work assembled in a digital edition. Funders should make detailed documentation a compulsory requirement. To help creators provide all of the necessary information, funders may wish to adopt a documentation template.

## Conclusion

This paper discusses the results of a web survey entitled "Expectations of Digital (Textual) Editions" circulated among the Digital Humanities academic community. The survey sought to give users of digital editions a platform to express their needs and expectations of digital editions of text. The survey ran for a month and recorded 218 completed responses, the highest response rate for a user survey recorded in the field to date. These user responses were compared against 242 digital editions collected by the *Catalogue of Digital Editions* project in order to identify meeting and diverging points between what creators of digital editions build and what users want. This comparative analysis assessed data both quantitatively and qualitatively to give the widest and most detailed picture to date of the two sides to digital (scholarly) editing, showing areas where resources can be better deployed, and user experience improved, by understanding the tools and features that a use community most desires, alongside those that have previously been delivered.

The results obtained from this study feed into previous studies on good practice in building Digital Humanities resources and crystallise the diverse range of needs of users of digital editions. The impression these results bestow upon the reader is that digital editions are imperfect tools unable to meet the expectations of every single user. While creators may feel discouraged by these results, one way to alleviate the negative sense of lenience exuding from these user responses might be to reconcile data reuse, licensing, image availability, and comprehensive documentation –the four most requested features– to the extent possible and to more clearly state motivations, objectives and intended audience.

To help better align digital editions to the needs of users and thus combat the risk of neglect, this study puts forward practical recommendations for both creators and funders of digital editions. These recommendations should not be considered mandatory but they conform to the needs expressed by the 218 participants of this survey.

The results of this research should not be ignored in the development of digital editions: it has signposted new avenues of enquiry along this line of research. For example, a future study could explore the extent to which creators of digital editions engage with their target users during the preparation and development stages of the project; another study might determine or categorise user needs according to user profiles (e.g. researcher versus general public, or versus student); another might shift its focus from the Digital Humanities to a wider user-base to address digital editions specifically designed and intended for other audiences; or, indeed, one may wish to create a survey specifically target at people who feel dissatisfied with existing digital editions. These and other possibilities speak to a field of research deserving of greater attention.

## Appendix A: Answers given to Question 19

The intention of this question was to draw out examples of digital editions that participants feel meet their needs. Some sixty digital editions were explicitly mentioned, and, as previously discussed, only eight of these were named by *Perseus Digital Library* (seven respondents), *Electronic Beowulf* (five respondents), *Folger Digital Texts* (four respondents), *Online Froissart* (three respondents), *Loeb Classical Library* (two respondents), *Corpus Corporum* (two respondents), *e-codices* (two respondents) and the *Bayeux Tapestry* (two respondents). Given that not all respondents who singled out these eight digital editions justified their answers, it is difficult at this stage to make an informed assessment as to the drivers behind their selection. *Perseus Digital Library*, *Folger Digital Texts*, *Loeb Classical Library* and



*Corpus Corporum* all provide a large number of texts, leading one to speculate that quantity plays a role in usage; similarly, *e-codices* provides high quality image reproductions of a large number of manuscripts; *Bayeux Tapestry*, *Online Froissart* and *Electronic Beowulf*, on the other hand, provide detailed contextual information, suggesting an appreciation for comprehensive introductions to the texts at hand.

Some respondents gave examples of digital editions that do not subscribe to the definition of digital edition used as reference in this survey (e.g. Response 64. *Free Courses On line* (Coursera, Open University, Stanford, MIT, etc. or Response 113. *Yellow 90s Online* (The Yellow Book)). These answers had no bearing over the analysis described in this paper but can be used for spin-off studies aimed at, for instance, better understanding digital editions of non-literary texts.

## Appendix B: Answers given to Question 20

As previously mentioned, although generally imposing a structure, Question 20 ('Is there anything else you would like to tell us about your user needs for digital editions that we have not covered here?') was designed to give respondents the opportunity to freely express their views. Comments given vary greatly: some corroborate the findings of the analysis of the digital editions collected in the Catalogue of Digital Editions, while others flag up issues or dimensions of digital editions that are currently overlooked, such as accessibility for people with disabilities, citation information and discoverability. The answers collected with this question are instrumental in the preparation of user-centric digital editions.

## Acknowledgments

The authors wish to thank all of the participants of the web survey for their key contribution toward a better understanding of users in digital (scholarly) editing. We also wish to thank our reviewers for their insightful comments and suggestions, as well as our collaborators Peter Andorfer and Ksenia Zaytseva from the Austrian Centre for Digital Humanities for their support in preparing this article.

## Bibliography

- Albers, Greg. "Bringing Books Online." *MW17: Museums and the Web 2017* (blog), 2017. <https://web.archive.org/web/20180220155910/https://mw17.mwconf.org/paper/the-next-generation-of-online-publishing-building-on-what-weve-learned-together/>.
- Allen, Elaine, and Christopher Seaman. "Statistics Roundtable: Likert Scales and Data Analyses." *Quality Progress*, no. July (2007). <http://rube.asq.org/quality-progress/2007/07/statistics/likert-scales-and-data-analyses.html>.
- Allison. "Other People's Data: Humanities Edition." *Journal of Cultural Analytics*, Debates, December 8, 2016. <https://web.archive.org/web/20170907081032/http://culturalanalytics.org/2016/12/other-peoples-data-humanities-edition/>.
- Apollon, Daniel, Claire B  lisle, and Philippe R  gnier, eds. *Digital Critical Editions: Exploring the Interweaving of Traditional and Digital Textual Scholarship*. Topics in the Digital Humanities. Chicago: University of Illinois Press, 2014. <http://www.press.uillinois.edu/books/catalog/92mby4hz9780252038402.html>.
- Bulatovic, Natasa, Timo Gnadt, Matteo Romanello, Juliane Stiller, and Klaus Thoden. "Usability in Digital Humanities - Evaluating User Interfaces, Infrastructural Components and the Use of Mobile Devices During Research Process." In *Research and Advanced Technology for Digital Libraries*, 335–46. Lecture Notes in Computer Science. Springer, Cham, 2016. doi:10.1007/978-3-319-43997-6\_26.
- Casadevall, Arturo, and Ferric C. Fang. "Reproducible Science." *Infection and Immunity* 78, no. 12 (December 1, 2010): 4972–75. doi:10.1128/IAI.00908-10.



- Cohen, Dan. "Open Access Publishing and Scholarly Values." *Dan Cohen*, May 27, 2010. <https://web.archive.org/web/20170906103757/http://www.dancohen.org/2010/05/27/open-access-publishing-and-scholarly-values/>.
- Dillen, Wout, and Vincent Neyt. "Digital Scholarly Editing within the Boundaries of Copyright Restrictions." *Digital Scholarship in the Humanities* 31, no. 4 (December 1, 2016): 785–96. doi:10.1093/llc/fqw011.
- Driscoll, Matthew James, and Elena Pierazzo, eds. *Digital Scholarly Editing: Theories and Practices*. Digital Humanities Series. Cambridge: Open Book Publishers, 2016. <http://books.openedition.org/obp/3381>.
- Fitzpatrick, Kathleen. "Open Access Publishing and Scholarly Values (Part Three)." *Planned Obsolescence*, May 28, 2010. <https://web.archive.org/web/20170906103634/http://www.plannedobsolescence.net/open-access-publishing-and-scholarly-values-part-three/>.
- Franzini, Greta. "Digital Critical Editions (2014) Daniel Apollon, Claire Bélisle and Philippe Régner." *Digital Scholarship in the Humanities* 30, no. 4 (December 1, 2015): 608–9. doi:10.1093/llc/fqv025.
- Franzini, Greta, Melissa Terras, and Simon Mahony. "A Catalogue of Digital Editions." In *Digital Scholarly Editing*, 1st ed., 4:161–82. Theories and Practices. Open Book Publishers, 2016. <http://www.jstor.org/stable/j.ctt1fzhh6v.13>.
- Galina, Isabel. "Creating a Regional DH Community – A Case Study of the RedHD." *Digital Humanities Quarterly* 9, no. 3 (2015). <http://www.digitalhumanities.org/dhq/vol/9/3/000221/000221.html>.
- Hamilton, Gill, and Fred Saunderson. *Open Licensing for Cultural Heritage*. Facet Publishing, 2017. [http://www.facetpublishing.co.uk/title.php?id=301850#.Wa\\_Ry9MjFE4](http://www.facetpublishing.co.uk/title.php?id=301850#.Wa_Ry9MjFE4).
- Hockey, Susan. *Electronic Texts in the Humanities: Principles and Practice*. Oxford: Oxford University Press, 2000.
- Horner, L. "Web Survey." Edited by Paul J. Lavrakas. *Encyclopedia of Survey Research Methods*. Sage Publications, 2011. doi:10.4135/9781412963947.n631.
- Hughes, Lorna M., ed. *Evaluating and Measuring the Value, Use and Impact of Digital Collections*. Facet Publishing, 2012.
- Kelly, Aodhán. "Tablet Computers for the Dissemination of Digital Scholarly Editions." *Manuscritica. Revista de Crítica Genética* 0, no. 28 (September 29, 2015): 123–40.
- Krippendorff, Klaus. *Content Analysis: An Introduction to Its Methodology*. SAGE, 2004.
- Lavagnino, John. "Access." *Literary and Linguistic Computing* 24, no. 1 (April 1, 2009): 63–76. doi:10.1093/llc/fqn038.
- McGann, Jerome. *A New Republic of Letters: Memory and Scholarship in the Age of Digital Reproduction*. Harvard University Press, 2014. <http://www.hup.harvard.edu/catalog.php?isbn=9780674728691>.
- Mierlo, Wim Van. "Electronic Textual Editing (Review)." *The Library: The Transactions of the Bibliographical Society* 9, no. 2 (July 17, 2008): 231–35.
- Montgomery, Scott L. *Does Science Need a Global Language? English and the Future of Research*. Chicago: University Of Chicago Press, 2013. <http://www.press.uchicago.edu/ucp/books/book/chicago/D/bo10984617.html>.
- Pierazzo, Elena. *Digital Scholarly Editing: Theories, Models and Methods*. Routledge, 2016.
- . "Quale Futuro per Le Edizioni Digitali? Dall'haute Couture Al Prêt-À-Porter." presented at the Fifth AIUCD Annual Conference, Venice, 2016. <http://www.himeros.eu/aiucd2016/a30.pdf>.
- Porter, Dorothy Carr. "Medievalists and the Scholarly Digital Edition." *Scholarly Editing* 34 (2013). [http://scholarlyediting.org/2013/essays/essay\\_porter.html](http://scholarlyediting.org/2013/essays/essay_porter.html).
- Price, Kenneth M. "Edition, Project, Database, Archive, Thematic Research Collection: What's in a Name?" 3, no. 3 (2009). <http://www.digitalhumanities.org/dhq/vol/3/3/000053/000053.html>.
- Ramsay, Stephen. "Open Access Publishing and Scholarly Values (Part Two)." *Stephen Ramsay*, May 28, 2010.

- <https://web.archive.org/web/20161105015022/http://stephenramsay.us/text/2010/05/28/open-access-publishing-and-scholarly-value-continued/>.
- Robinson, Peter. "Electronic Editions for Everyone." In *Text and Genre in Reconstruction : Effects of Digitalization on Ideas, Behaviours, Products and Institutions*, by Willard McCarty, 145–63. Digital Humanities Series. Cambridge: Open Book Publishers, 2010. <http://books.openedition.org/obp/656>.
- . "The Canterbury Tales and Other Medieval Texts." In *Electronic Textual Editing*, edited by Lou Burnard, Katherine O'Brien O'Keffe, and John Unsworth. Modern Language Association of America, 2006. [http://www.tei-c.org/About/Archive\\_new/ETE/Preview/robinson.xml](http://www.tei-c.org/About/Archive_new/ETE/Preview/robinson.xml).
- Sahle, Patrick. "About." Text. *A Catalog of Digital Scholarly Editions*, April 1, 2008. <http://www.digitale-edition.de/vlet-about.html>.
- . "Kriterien Für Die Besprechung Digitaler Editionen, Version 1.0 |." Institut für Dokumentologie und Editorik, 2014. <https://www.i-de.de/publikationen/weitereschriften/kriterien-version-1-1/>.
- Schmidt, Desmond. "Towards an Interoperable Digital Scholarly Edition." *Journal of the Text Encoding Initiative*, no. Issue 7 (November 12, 2014). doi:10.4000/jtei.979.
- Schouten, Barry, Fannie Cobben, and Jelke Bethlehem. "Indicators for the Representativeness of Survey Response." *Survey Methodology* 35, no. 1 (2009): 101–13.
- Seidlhofer, Barbara. *Understanding English as a Lingua Franca - Oxford Applied Linguistics*. Oxford University Press, 2013.
- Sutherland, Professor Kathryn, and Professor Marilyn Deegan. *Text Editing, Print and the Digital World*. Ashgate Publishing, Ltd., 2012.
- Terras, Melissa. "Quantifying Digital Humanities," 2011. <https://www.flickr.com/photos/ucldh/6730021199/sizes/o/in/photostream/>.
- . "Should We Just Send a Copy? Digitisation, Usefulness and Users." Accessed September 2, 2017. <http://connection.ebscohost.com/c/articles/49547304/should-we-just-send-copy-digitisation-usefulness-users>.
- Tomasi, Francesca. "Digital editions as a new model of conceptual authority data." *JLIS.it* 4, no. 2 (2013): 21–44.
- Van Hulle, Dirk. "Annotated Bibliography: Key Works in the Theory of Textual Editing." In *Electronic Textual Editing*, edited by Lou Burnard, Katherine O'Brien O'Keffe, and John Unsworth. Modern Language Association of America, 2006. [http://www.tei-c.org/About/Archive\\_new/ETE/Preview/vanh-bib.xml](http://www.tei-c.org/About/Archive_new/ETE/Preview/vanh-bib.xml).
- Vanhoutte, Edward. "Defining Electronic Editions: A Historical and Functional Perspective." In *Text and Genre in Reconstruction: Effects of Digitalization on Ideas, Behaviours, Products and Institutions*, edited by Willard McCarty, 119–44. Digital Humanities. Cambridge: Open Book Publishers, 2010. <http://books.openedition.org/obp/654#notes>.
- Visconti, Amanda. "An Invitation to Beta-Test the Infinite Ulysses Digital Edition - Maryland Institute for Technology in the Humanities," 2015. <https://web.archive.org/web/20170906103921/http://mith.umd.edu/invitation-beta-test-infinite-ulysses-digital-edition/>.
- Warwick, Claire. "Studying Users in Digital Humanities." In *Digital Humanities in Practice*, edited by Claire Warwick, Melissa Terras, and Julianne Nyhan, 1–21. Facet Publishing, 2012.
- Warwick, Claire, Isabel Galina, Melissa Terras, Paul Huntington, and Nikoleta Pappa. "The Master Builders: LAIRAH Research on Good Practice in the Construction of Digital Humanities Projects." *Literary and Linguistic Computing* 23, no. 3 (September 1, 2008): 383–96. doi:10.1093/lc/fqn017.
- Warwick, Claire, Melissa Terras, Isabel Galina, Paul Huntington, and Nikoleta Pappa. "The Master Builders: LAIRAH Research on Good Practice in the Construction of Digital Humanities Projects." In *Digital Humanites 2007: The 19th Joint International Conference of the Association for Computing in the Humanities and the Association for Literary and Linguistic Computing*. University of Illinois, Urbana Champaign, June 4-8, 2007. Urbana Champaign, 2007. <http://discovery.ucl.ac.uk/4807/>.