Towards a first digital edition of the oldest surviving manuscript of St Augustine’s *De civitate Dei*

*Greta Hayley Franzini*

A dissertation submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

of

University College London.

Department of Information Studies

University College London

24th April 2019
I, Greta Hayley Franzini, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the work.

Date ....................... Signature.................................................................
Abstract

This thesis describes the creation of a pilot digital edition of MS XXVIII(26), the oldest surviving manuscript of Saint Augustine’s (354-430 AD) monumental *De civitate Dei* (The City of God). Also known as Manuscript V[eronensis], MS XXVIII(26) dates back to the early fifth century AD and is housed in the chapter library of Verona, Italy. As contemporary to Saint Augustine himself, it is a particularly treasured object of study. This thesis reassesses extant research about this manuscript, collecting information about its disputed provenance, historical context, materiality, tradition, and conservation. In doing so, it investigates how the manuscript can be best reproduced as a digital edition by way of two surveys designed to better understand how digital editions are respectively being created and used.

The survey devoted to the study of how digital editions are being *built* has become a publicly available digital resource in collaboration with the Austrian Academy of Sciences. The resource, known as the *Catalogue of Digital Editions*, aggregates and catalogues a large number of digital editions in an effort to delineate the field’s *status quo* and spawn new quantitative and qualitative research.

The community survey devoted to the study of how digital editions are being *used* is one of the very few as well as the largest in the field yet. The over 200 responses received give detailed information regarding the expectations of digital editions provided by the Digital Humanities community and point to many areas for further improvement.
Abstract

A comparative analysis of the results from the two surveys suggests that while creators are aware of and adhere to standards of creation, much work remains to be done to address the needs of a diverse range of users. With this information, digital editors in the Digital Humanities can better shape future projects and thus contribute to the production of ever-useful digital cultural resources. This information is also guiding the creation of a pilot digital edition of MS XXVIII(26), which remains to be user-tested but serves as the first digital reproduction of the oldest surviving manuscript of Saint Augustine’s *De civitate Dei*.

The research described in this thesis has led to the formulation of recommendations for those embarking on the creation of a digital edition. Specifically, creators are advised to get access to the original documents and to high resolution images, to provide transcriptions of the text in multiple formats so as to enable further research and data reuse in a variety of academic contexts, to provide detailed documentation of the editorial and technological components of the project, to make as much data available under open licences and, finally, to conduct, and report on, user studies of the digital edition.
Impact Statement

Within academia, the relevance of the comparative analysis that grounds this thesis was recognised early on and led to the formalisation of a collaboration between myself and the Austrian Centre for Digital Humanities at the Austrian Academy of Sciences in Vienna to offer a useful means of studying, evaluating and (re)using the data collected for new quantitative and qualitative research. The digital resource jointly developed allows users to browse, visualise and filter the data around their research interests. Its integration with relevant Linked Open Data vocabularies enriches, organises and connects the data to the wider semantic web, traversing discipline and domain boundaries; supporting granular interrogation and knowledge discovery; aggregating distributed resources to reduce dispersion; and structuring data for improved search engine indexing, thus increasing the visibility and discoverability of the projects collected. This thesis also contributes the first overarching detailed comparison in the Digital Humanities domain of user requirements and expectations of digital editions against the digital editions being built, consequently delineating the field’s status quo. As well as highlighting meeting and diverging points between creation and use, this comparative analysis might inform the development of digital editions better tuned to both individual and community expectations, concurrently optimising existing methods or shaping new avenues of enquiry. The results of this analysis have been accepted for publication in a special issue on evaluation of digital resources of the Association for Computing Machinery’s Journal on Computing and Cultural Heritage.
The resource has been accessed by thousands of people across six continents (Asia, Africa, North America, South America, Europe and Australia); it is referenced and cited in educational resources (e.g., #dariahTeach MOOC Scholarly Digital Editions: Manuscripts, Texts and TEI Encoding), university pages and publications as a key resource for those studying digital scholarly editing; it receives external submissions from users, who wish to disseminate their efforts and that of their country; it was nominated and won third place in the Digital Humanities 2016 Awards; it is syndicated by the German Library Network of Scientific Databases (DBIS), attesting to its value to libraries and hence penetrating the non-academic world. Whether accessing the project directly or via the more than 300 libraries subscribing to the DBIS service, users from all backgrounds, both within and outside academia, discover digital editions, the cultural heritage these showcase, the institutions involved in their creation and related initiatives. Most recently, the project was broadcast on Italian national television (RAI Cultura and RAI Scuola) as part of an interview and will feature in an upcoming educational series about computational processing of cultural heritage for the general public.

The results of this research have been presented internationally at conferences (i.e., Digital Humanities 2013 conference), seminar series (i.e., Digital Classicist London and Digital Humanities Seminar Series of the Fondazione Bruno Kessler), expert workshops (i.e., NeDiMAH, DARIAH-DE), and in two publications. Specifically, a book chapter in the open access volume Digital Scholarly Editing: Theories and Practices describes an early version of the resource (Franzini et al.,

---

2016), while a journal article accepted for publication in a special issue on evaluation of digital cultural resources of the *Journal on Computing and Cultural Heritage* examines the relation between creation and use of digital editions (Franzini et al., 2019).

The impact of this research is expected to continuously and incrementally grow in the coming years with the further development of the published resource, its dissemination through scholarly articles, conference participation and social media, and through the pursuit of new collaborations with GLAM institutions.
Acknowledgements

I would like to express my eternal gratitude to my advisers Melissa Terras and Simon Mahony for the continuous support, patience, care and encouragement throughout this journey. Their guidance helped me become a better researcher and writer. I could not have wished for better companions.

I dedicate this thesis to my beautiful family, my perpetual source of strength and happiness.
Contents

1 Introduction 31
  1.1 Premise . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 31
  1.2 Thesis structure . . . . . . . . . . . . . . . . . . . . . . . . . . . 32

2 Digital editions 35
  2.1 A short history . . . . . . . . . . . . . . . . . . . . . . . . . . . . 36
  2.2 What is an edition? . . . . . . . . . . . . . . . . . . . . . . . . . 40
  2.3 What is a digital edition? . . . . . . . . . . . . . . . . . . . . . . 46
  2.4 Advantages of digital editions . . . . . . . . . . . . . . . . . . . . 48
    2.4.1 Infrastructure for digital editions . . . . . . . . . . . . . . 53
    2.4.2 Challenges of digital editions . . . . . . . . . . . . . . . . 56
    2.4.3 The status quo . . . . . . . . . . . . . . . . . . . . . . . . 59

3 A Catalogue of Digital Editions 61
  3.1 Why build a Catalogue of Digital Editions? . . . . . . . . . . . . 62
  3.2 Method . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 63
    3.2.1 Building the Catalogue of Digital Editions . . . . . . . . . 64
    3.2.2 Content analysis: data collection and evaluation . . . . . 70
  3.3 The Web Application . . . . . . . . . . . . . . . . . . . . . . . . . 74
    3.3.1 Architecture and data visualisations . . . . . . . . . . . . 74
    3.3.2 Dissemination . . . . . . . . . . . . . . . . . . . . . . . . . 91
    3.3.3 User analytics . . . . . . . . . . . . . . . . . . . . . . . . . 92
  3.4 Sustainability . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 95
Contents

3.5 Future work ........................................ 96
3.6 Conclusion ........................................ 96

4 User expectations of digital editions ........................................ 99
4.1 Users, use, usage, usefulness and usability ......................... 100
  4.1.1 Users ........................................ 101
  4.1.2 Use .......................................... 102
  4.1.3 Usefulness .................................... 102
  4.1.4 Usability ..................................... 103
4.2 Related work: Users studies of digital editions .................... 104
4.3 Method ................................................ 109
  4.3.1 Survey design .................................. 110
4.4 Discussion and recommendations ....................................... 145
4.5 Conclusion ........................................... 147

5 Reassessing the oldest surviving manuscript of De civitate Dei: MS XXVIII(26), Biblioteca Capitolare di Verona ........................................ 153
  5.1 Verona in the 4th and 5th centuries AD .............................. 154
  5.2 The Biblioteca Capitolare .................................. 156
  5.3 St Augustine and his De civitate Dei .............................. 157
    5.3.1 The tradition of De civitate Dei ..................... 160
  5.4 MS XXVIII(26) ....................................... 161
    5.4.1 Previous research .................................. 161
    5.4.2 The manuscript .................................... 164
    5.4.3 Conservation status ................................. 180
    5.4.4 Comparing MS XXVIII(26) to coeval manuscripts .... 181
  5.5 Why is MS XXVIII(26) a good candidate for a digital edition? .. 189
  5.6 Conclusion ........................................... 190

6 A pilot, semi-diplomatic documentary digital edition of MS XXVIII(26) ........................................ 193
  6.1 Methodology ........................................... 194
    6.1.1 Transcription ...................................... 197
Contents

7.5.4 Training ................................................. 229
7.5.5 Users .................................................. 229
7.5.6 Dissemination ......................................... 229

Bibliography .................................................. 232

Appendices .................................................... 252

A ‘Expectations of Digital (Textual) Editions’ Web Survey .................. 253
A.1 Survey email announcement .................................. 253
A.2 Answers given to Question 3 ................................. 254
A.3 Answers given to Question 19 ............................... 260
A.4 Answers given to Question 20 ............................... 266
A.5 Full survey report: complete and incomplete responses. ........ 270

B Chronology: St Augustine and MS XXVIII(26) ......................... 303

C Conservation report of MS XXVIII(26) .......................... 307
List of Figures


2.2 A work by Cicero as edited in the *Oxford Classical Text* series. Numbers in the margin are used to mark sections and lines, while the bottom part of the page accommodates the critical apparatus. Source: https://archive.org/stream/CiceroDeRePublicaOxfordZetzel/Cicero%2C%20De%20Re%20Publica%20%28Oxford,%20Zetzel%29#page/n249/mode/2up (Accessed: 27 January 2018). 42

2.3 A work by Sulpicius Severus as edited in the *Corpus Scriptorum Ecclesiasticorum Latinorum* edition of Latin Church fathers. Here, numbers in the margin mark sections and lines, while at the bottom of the page space is given for notes (the standalone line string of text placed between the main text and the critical apparatus) and to the critical apparatus. Source: http://archive.org/stream/corpusscriptoru12wissgoog#page/n31/mode/2up (Accessed: 27 January 2018). 43

2.4 A 1709 edition of a work by Greek sophist Philostratus divides the page into three main areas: one column for the Greek text, one column for the Latin translation, and a bottom section for the critical apparatus; the order of the columns is alternated between pages. Source: http://archive.org/stream/tatonphilostrato00phil/page/302/mode/2up 44
2.5 The *Chopin’s First Editions Online* project ([http://www.chopinonline.ac.uk/cfeo/](http://www.chopinonline.ac.uk/cfeo/)) allows users to choose editions to compare; the images can be enlarged for a detailed study of the similarities and differences. This screenshot shows a comparison of two editions of *Rondo pour le Pianoforte, Op.1.*

2.6 The Great Isaiah Scroll of the *Digital Dead Sea Scrolls* project can be (un)rolled and read in a dynamic fashion using the text-image link and macro-view features. Source: [http://dss.collections.imj.org.il/isaiah#2:4](http://dss.collections.imj.org.il/isaiah#2:4) (Accessed: 27 January 2018).

2.7 *The Digital Vercelli Book* project makes use of the *Edition Visualization Technology* (EVT) to publish XML-encoded text and manuscript images in the web in a functional reading environment.

3.1 GitHub repository of the Catalogue of Digital Editions (September 2017 snapshot).

3.2 Repositories created with free GitHub accounts come with seven information layers (September 2017 snapshot).

3.3 The Code layer of the Catalogue’s GitHub repository, containing all the files that make up the Catalogue (April 2017 snapshot).

3.4 GitHub’s versioning system allows users to track all changes made to a file (April 2017 snapshot).

3.5 The Issues page of the Catalogue lists issues and tasks (September 2017 snapshot).

3.6 The landing page of the Catalogue’s GitHub Wiki (September 2017 snapshot).

3.7 The `digEds_cat.csv` file as displayed by GitHub. The file can be searched using the search bar provided and a scrolling option allows users to move across the spreadsheet both vertically and horizontally (April 2017 snapshot).

3.8 Example digital edition in the Catalogue as displayed by the GitHub .csv view.
3.9 Logo of the Catalogue of Digital Editions (my design). . . . . . . 75
3.10 Homepage of the web application of the Catalogue of Digital Edi-
tions (April 2017 snapshot). . . . . . . . . . . . . . . . . . . . . . . . 76
3.11 The web application’s Browse the Catalogue page. Users can
either serendipitously browse the digital editions or look for spe-
cific projects by using the faceted search provided (April 2017
snapshot). . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 77
3.12 Users who are not looking for a particular digital edition can
browse the Catalogue with the help of a pagination function at
the bottom of the screen, allowing them to jump between pages
of data (April 2017 snapshot). . . . . . . . . . . . . . . . . . . . . . . . . . 77
3.13 Example digital edition sheet in the Catalogue, with project de-
tails on the left and a map pointing to the involved institutions
on the right. Clicking on the location marker brings up a pop-up
window with more information about the institution (April 2017
snapshot). . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 79
3.14 The i information button, when pressed, displays the descrip-
tions of each feature in a collapsible column placed between the
feature column and the value column (second column in this im-
age) (April 2017 snapshot). . . . . . . . . . . . . . . . . . . . . . . . . . 79
3.15 A share button at the bottom of each digital edition page in the
Catalogue can be used to share a project via Twitter, Facebook
and Google Plus. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 80
3.16 A red banner above the sheet header of the Saint-Aubins Book of
Arses digital edition informs users that the project is no longer
available under the URL originally provided (April 2017 snapshot). 80
3.17 The user can choose one of three types of visualisation (bar, pie
or line chart) for the list of features provided (April 2017 snapshot). 81
3.18 A menu icon in the top right corner of every Catalogue chart
opens up download options in a pop-up window. . . . . . . . . . . . . 81
3.19 Two different views, a bar and a pie chart, of the number of digital editions produced per country. Hovering over the Italian slice of the pie brings up a pop-up window with the exact count of digital editions produced in Italy (April 2017 snapshot).  

3.20 Google Chrome’s web translation plugin automatically detects the language of a particular page and offers the user the option to translate the text into another language. The plugin relies on the machine translation algorithms of Google Translate.  

3.21 Coverage of historical periods in the Catalogue (April 2017 snapshot).  

3.22 Almost 150 digital editions in the Catalogue do not make their contents available in open access (April 2017 snapshot).  

3.23 The largest group of digital editions in the Catalogue encode their electronic texts in TEI-compliant XML (April 2017 snapshot).  

3.24 The majority of digital editions currently in the Catalogue provide digital images of the primary source(s) (April 2017 snapshot).  

3.25 Map visualisation of the institutions recorded in the Catalogue. The only marker details a single digital edition from the Indian Ministry of Culture; when clicking on the digital edition, the user is directed to the project’s sheet in the Catalogue (April 2017 snapshot).  

3.26 A user request made via the web application’s GitHub repository asked for the provision of permalinks for all projects in the Catalogue. While the application currently provides URIs, these are not stable identifiers yet.  

3.27 A removable blue information box explains how to use previous/next arrows to navigate the Catalogue (April 2017 snapshot).
3.28 The Catalogue’s Piwik analytics landing page. The primary menu is located in the leftmost side-bar and the majority of screen space is given to the data widgets (September 2017 snapshot).

3.29 Data widgets can be customised to view specific layers of information (September 2017 snapshot).

3.30 The Piwik Visitor Log lists visitors of the Catalogue. Moving the cursor over the three middle icons will bring up more information in a black pop-up window. Alternatively, the user can click on View visitor profile in the top right corner to enlarge the visitor card (April 2017 snapshot).

3.31 On 18 April 2017 Piwik revealed that the Catalogue had been included in the German Datenbank-Infosystem (DBIS) (April 2017 snapshot).

3.32 The Catalogue record as shown by the Datenbank-Infosystem (DBIS) (April 2017 snapshot).

4.1 The chosen software (see further below) anonymously lists all participants of the survey, assigns them a unique ID and details the time-stamp and duration of their participation.

4.2 Answers recorded for survey question 1.

4.3 Answers recorded for survey question 2.

4.4 Answers recorded for survey question 4.

4.5 Answers recorded for survey question 5.

4.6 Answers recorded for survey question 6.

4.7 Answers recorded for survey question 7.

4.8 Answers recorded for survey question 8.

4.9 Answers recorded for survey question 9.

4.10 Answers recorded for survey question 10.

4.11 Answers recorded for survey question 11.

4.12 Answers recorded for survey question 12.

4.13 Answers recorded for survey question 13.
4.14 Answers recorded for survey question 14. ............................................. 136
4.15 Answers recorded for survey question 15. ............................................. 138
4.16 Answers recorded for survey question 16. ............................................. 139
4.17 Answers recorded for survey question 17. ............................................. 149
4.18 Answers recorded for survey question 18. ............................................. 150
4.19 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. ................................................................. 151
5.1 Zoomed-in view of Ancient Northern Italy. Verona (in the centre) is located at the intersection of three arterial roads: the Via Postumia (in light blue), built by the Romans in 148 BC; the Via Gallia, which provided a connection with Brixia (Brescia) and Mediolanum (Milano) (thin black); and the Via Claudia Augusta (thick black), leading up to modern day Austria and Germany via Tridentum (Trento). Zoomed version of Wikimedia Commons image available at: http://commons.wikimedia.org/wiki/File:Viae_Flaminia_Aemilia_Postumia.jpg (Accessed: 8 October 2014). .......... 156
5.2 Jones’ breakdown of Augustine’s publication strategy. ......................... 159
5.3 Dombart’s stemma codicum of De civitate Dei (Dombart, 1908, pp. 33-34). Roman and Greek letters are used to represent manuscripts, Ω (= omega) being the archetype or first text; edges are used to visualise the relationships between the manuscripts and the timeline on the left hand-side, starting at the top with the fifth and ending with the fifteenth century, adds a temporal dimension to the stemma. ................................................................. 161
5.4 Screenshot of an ORBIS map displaying the fastest (red), cheapest (green) and shortest (purple) summer routes from Hippo Regius to Verona. ORBIS is available at: http://orbis.stanford.edu/ (Accessed: 17 October 2014).

5.5 From top to bottom. The top row represents the two-volume format envisaged by Augustine for *De civitate Dei*. Here, volume one contains ten books and volume two contains twelve books; in this two-volume format, MS XXVIII(26), depicted in red, would be enclosed in volume two. Conversely, the lowest row of the table represents the five-volume format of *De civitate Dei* as envisaged by Augustine; here, MS XXVIII(26) would enclose volume three and the first two books of volume four. Table my own creation.

5.6 Follicle grouping and pattern of folio 50r. The triplet grouping is common to goatskin. Pattern colouring my own.

5.7 Ruling marks, folio 43v: line 18 (above) and lines 27 and 28 (below). Composition my own.

5.8 The front matter of MS XXVIII(26) reads: Beat(us) aug(ustinus) de ci(ui)tate dei l(ibri) VI uid(e)l(ic)et XI, XII, XIII, XIII, XV, XVI.

5.9 *AE* diphthong represented by a minuscule letter *E* with descender, also known as *e caudata* (folio 1v, line 7).

5.10 Carolingian and insular minuscule letter *G* (folio 1v, line 28).

5.11 The *ET* ligature looks like an ampersand (folio 1r, line 25). The symbol we use today as a replacement for the English conjunction *AND* is in fact a combination of the uncial characters *e* and *t*, which make up the Latin conjunction *et* (= *and*). In MS XXVIII(26) this ampersand symbol is also inconsistently used to replace *-et* verb endings or any other occurrence of this letter pairing.
5.12 The minuscule letter r is combined with an especially long i (folio 2r, line 4). ........................................ 172
5.13 Two types of decorative markers, both labelling the beginning of a new chapter: on the left (folio 2v) the marker appears to be in the shape of a leaf; on the right (folio 12v), the capital letter C and the Roman numeral seven indicate Capitulum seven starts with the word UIDEMUS. ........................................ 173
5.14 The title on folio 1v contains decorative motifs acting as macra. LIB is an abbreviation of LIBRI and DI of DEI. .......................... 173
5.15 Characteristic uncial letters D, E, A and M in line 13, folio 73v. ... 175
5.16 Minuscule letters Q and H in line 14, folio 57r. ......................... 175
5.17 Ink corrosion on folio 74r (lines 27-30). The different inks are clearly visible: the original gall-ink is visible at the beginning of sentence 30 with the first half of the word corporales (CORP-); the second half of the word (-ORALES) and, indeed, the rest of the sentence was retraced and presents noticeable signs of corrosion. 176
5.18 Part of the text in folio 61r is completely hidden by mould. Even the scholia (blue focus box) are barely visible. ......................... 176
5.19 The halo surrounding the abbreviated word INCIPIT on folio 76v is caused by the reaction of the mould against the iron gall ink. .. 177
5.20 This Christogram, drawn on folio 7r, appears to be an original feature (fifth century). ................................................... 178
5.21 Folio 252v dates back to the eighth century and contains a Catholic liturgy. ......................................................... 179
5.22 Codex Bobbiensis (k) (5th cent.). African uncial. Source:
http://en.wikipedia.org/wiki/Codex_Bobiensis#mediaviewer/File:
5.23 Codex Veronensis - Evangelarium Purpureum (4-5th cent.). Italian uncial. Source: http://en.wikipedia.org/wiki/Codex_Veronensis#
5.24 MS XXVIII(26), extract from folio 57v. 186


6.1 An example of ligature encoding. 200

6.2 The <app> provides the readings of manuscript BSB-Hss Clm 6267 (#Clm6267 in the XML), of Codex Sangallensis 178 (#Sang178) and of the standard CSEL edition (#CSEL40) in a single @wit attribute. 201

6.3 The macron’s location in this example (word ENIM in folio 12v, line 8) is preserved by adding a space between the letter ‘I’ and the macron HTML character code. 201

6.4 Example of how the <choice> tag is used to encode the abbreviated word ENIM. 202

6.5 The letter ‘O’ in the word ABSORDA has been deleted and replaced with a ‘U’ (folio 25r, line 13). 202

6.6 The word TERRAM below is missing one ‘R’ and is, therefore, tagged as an error (folio 14v, line 25). 202

6.7 The word HABENT in folio 18r is split between lines 22 and 23; <w> tags are thus used to denote the word break. 203

6.8 Example of damaged text from folio 13r, line 16, and its corresponding encoding. 203

6.9 Scribal addition of PRO between UT and EO (folio 35r, line 9) appears above the main text and is thus encoded accordingly. 203

6.10 Example of expunction in folio 10v, line 27. 204

6.11 Scribal deletion and encoding of the letter ‘A’ in SEPTIMOQUAE (folio 12v, line 21). 204

6.12 The pronoun EIUUS is followed by what was once a second EIUUS, which was subsequently turned into the verb EST by erasing the letters ‘I’ and ‘U’ and by adding a ‘T’ (folio 38r, line 6). 204
26

List of Figures
6.13 The incorrect ‘S’ at the end of the word UULGATISSIMA has been
overlooked. While the editor does not explicitly identify the mistake, the <app> element implicitly reveals the correct reading. . .

205

6.14 The encoded omission of SENARIAM UEL. . . . . . . . . . . . . .

205

6.15 Illegible annotation in folio 13r, around line 4, and its respective
encoding. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .

206

6.16 This leaf signature at the bottom of folio 38v stands for q(uaterni)
IIII (Roman numeral for 4, often also written as IV ), which tells
the reader that this is the fourth quire in the codex. Signatures of
quires one, two and three can be seen on folia 14v, 22v and 30v
respectively. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .

207

6.17 An encoding solution for word breaks. . . . . . . . . . . . . . . .

207

6.18 Snippet from folia 29r and 29v. . . . . . . . . . . . . . . . . . . . .

208

6.19 The first string of XSL code in the top of the screenshot is used
to tell the XSL parser to retrieve information from the <note>
element in the XML files; the second string of XSL code is used
to tell the parser to retrieve the number of the chapter stored in
the @n attribute of the <note> XML tag. The final HTML result,
Chapter 1, is visible in the bottom part of the screenshot. . . . . .

210

6.20 The file system underlying the digital edition of MS XXVIII(26)
stores all files making up the capitula, back matter and individual
books in separate folders for better . . . . . . . . . . . . . . . . . .

211


212

6.22 A top navigation menu lists the eight textual divisions in the manuscript; clicking on any one opens a drop-down list of all folia
making up a particular division. Alternatively, users can use the
bottom ‘Previous/Next’ pagination for a sequential reading of the
text. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .

213


6.23 The two-tab layout of the textual region of the edition page provides both the transcribed text and automatically generated notes listing variants encountered in other witnesses. 214

6.25 Default CSS and JavaScript Bootstrap settings must be added to the <head> of the website’s HTML documents. Customisations are possible and must be added as separate files in the <head>. 214

6.24 JavaScript code is used to give users the possibility of enlarging any portion of an image. 215

6.26 Scribal additions (in green) and editorial additions (in cerise), folio 7r. 215

6.27 To visualise variant readings of a text, the TRAViz user need only add plain text transcriptions of the manuscript witnesses to the HTML file provided; in this example, I have added five different texts to compare. 219

6.28 Excerpt of book 11, chapter 3 of De civitate Dei (folio 9r, lines 8-9 of MS XXVIII(26)). For every word in a graph a pop-window lists all witnesses in which the word appears. 220
List of Tables

5.1  List of abbreviations occurring in MS XXVIII(26) with the corresponding expansions.  178

5.2  Manuscripts contemporary to MS XXVIII(26).  182

6.1  Unicode characters used in the transcription of MS XXVIII(26) to reproduce special characters, abbreviations and scribal conventions.  199
Chapter 1

Introduction

1.1 Premise

Recent years have witnessed a surge in the production of digital editions of manuscript culture. The growing commitment of libraries, museums and other cultural memory institutions to digitise their holdings incites the creation of computational solutions for information enrichment, knowledge discovery and dissemination. One such solution, the digital scholarly edition, stays true to established philological methods of enquiry but its innovative spirit and technological flexibility beats the obstacles posed by its printed counterpart, giving creators the opportunity to study and share manuscript culture at a global scale in previously impossible ways.

Operating at the intersection of humanistic and information sciences, in a field that can be defined as Digital Humanities, digital editions prove to be invaluable instruments to advance research, and to increase the visibility and public memory of codices, dispersed manuscript pages, correspondence, diaries, notebooks, atlases, books, newspapers, charters, inscription or tablets. The beginning of digital editing practice can be traced back to the 1980s and, since then, dozens of research publications have discussed the theories, methods and impact of digital scholarly editions, modelling them upon their print counterparts while pushing the boundaries towards original and experimental means of representa-
Yet, despite the headway made by the community of digital editing practitioners in outlining the theories, methods and models underpinning digital editions, little is known about how these translate into practice. In other words, the amount of theoretical literature published outnumbers the number of reports documenting the practical implementation of these notions. Similarly, while much is known about the creation of digital scholarly editions, the same cannot be said about the use of these resources, which have received little to no attention. This thesis addresses both of these gaps. Firstly, it charts and examines a large number of digital editions in a catalogue to better define the status quo in the field and to provide evidence upon which we can further develop our understanding of the practice of digital scholarly editing; the data collected can help both users and creators of digital editions assess efforts and identify patterns of creation, cross-correlating to the theories of scholarly editing. This catalogue seeks to become a gateway to vast realm of digital editions in support of knowledge discovery and to help shape future projects. Secondly, this thesis presents the results of a web survey on the use and expectations of digital scholarly editions, and compares the findings to the data collected in the catalogue. In doing so, this work provides a first understanding of the overlap between creation and use of these digital cultural resources.

1.2 Thesis structure

There are two thematic parts to this thesis. The first part addresses the field of digital editing as a whole. Chapter 2 summarises the field, outlining the history of (digital) editions and the advantages born by technology in the digital reproduction of manuscript culture. Chapter 3 describes the motivation behind, and the creation of, the Catalogue of Digital Editions, a project that started as a personal tool of study but that has taken on a life of its own and is now a successful web application and collaboration with the Austrian Centre for Digital Human-
1.2. Thesis structure

The Catalogue of Digital Editions is working towards the adoption of library cataloguing standards in an effort to advocate the value of digital editions as bibliographic records. The syndication of the Catalogue of Digital Editions in the German Network of Online Databanks (DBIS) is the first decisive step in this direction as digital editions present in the Catalogue are now available to over 300 libraries in Germany. Chapter 4 describes a survey entitled "Expectations of Digital (Textual) Editions" designed to better understand what users expect or want from a digital edition. The first survey of its kind, its response rate is unprecedented, with 218 complete questionnaires; the correlation of responses to the data collected in the Catalogue proffered practical recommendations for creators and funders of digital editions upon which new research can be formulated to increase the use and usefulness of these complex resources.

The second part of this thesis moves away from general considerations to take a closer look at my proof-of-concept digital edition of the oldest surviving manuscript of St Augustine’s magnum opus De civitate Dei. More specifically, Chapter 5 reassesses the manuscript, known as MS XXVIII(26) or Manuscript V, by bringing together all extant literature and provides the only conservation report to have ever been carried out on this manuscript (by Dr Alberto Campagnolo). Chapter 6 describes the production of the proof-of-concept digital edition and discusses how this pilot digital edition reconciles the needs of users identified in Chapter 4 with my own interests, outlining future work.

The general conclusions sum up the original contribution of this thesis and give recommendations to researchers and institutions, reflecting on the open challenges in the field of digital editing in general and of this research in particular.
Chapter 2

Digital editions

A digital 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. The growing commitment of libraries, museums and other cultural memory institutions to digitise their holdings requires the provision of information to accompany the millions of images now accessible through the web. The creative and technological flexibility afforded by digital editions today broadens their potential, giving researchers the opportunity to explore a wide variety of documents in ever-greater detail. Operating at the intersection of humanities and information science, digital editions are valuable instruments for the advancement of research and knowledge, and for the dissemination, accessibility, public engagement and public memory of the world’s manuscript culture.

This first chapter introduces editions and digital editions with a view to understanding their origin, purpose, structure and use. It does not provide an extensive history of textual scholarship and digital scholarly editing, as that has already been thoroughly covered by Vanhoutte (2010) and Greetham (2013). While relevant to the development of any research question on digital scholarly editing, the theory of editing, and the numerous debates had over the years to tackle questions such as “what is a text?” (Hockey et al., 1999) or “what is the editor” (Vanhoutte, 2009), are not central to this thesis, which, rather, deals with
data and with the information we can derive about both the creation and use of
digital editions.

2.1 A short history

The field of (digital) scholarly editing spawns from Textual Studies, a historical
field of scholarship, whose origins can be traced as far back as Antiquity (Turner,
2014).¹ Textual Studies concerns itself with reading and writing culture, and de-
velops across many disciplines, each addressing a particular aspect of a text. One
of these is Textual Criticism, a term often used interchangeably with the more
ambiguous Philology, which:

 designate a set of techniques or operations whose purpose is to re-
construct texts whose complete genuineness is open to doubt; which
is to say practically all texts, in any ancient or modern language, that
have not come down to us in a faultless autograph or in a copy that
the author —sometimes a poor copyist of his own works— carefully
revised, and especially texts transmitted in a series of copies that are
more or less far removed from the original, that is to say, deformed
by a number of errors. (Trovato, 2014, p. 39)

These techniques include Stemmatology, which concerns itself with the definition
of the genealogy of a text; Palaeography, the study of historical handwriting; or
Codicology, which is the study of manuscripts and codices as cultural artifacts.
Other disciplines that fall within the umbrella of Textual Studies are Literary Cri-
ticism, which is the study of literature; Genetic Criticism, the study of the chro-
nological or sequential development of a text or exemplar; but also Diplomatics,
text can encapsulate any number of these techniques and disciplines depending

¹Published volumes on Textual Studies or Scholarship include Variants - The Journal of the
European Society for Textual Scholarship, available at: http://journals.openedition.org/
and Flanders (2013), Modiano et al. (2017).
on its rationale and purpose.

To write about the history of (digital) editions is to travel some two thousand years back in time. The poet Antimachus of Colophon (fifth-fourth centuries BC) is known to have produced the first edition of the Homeric poems, while the first *standard* editions (those based on a *collation* or comparison of manuscripts) of the Homeric tradition date back to the Alexandrian period, more specifically to 275 BC with the Greek philologist Zenodotus (Turner, 2014, pp. 9-10). The Derveni Papyrus, the oldest literary papyrus to have been discovered, has been dated to 340-320 BC and provides, among other texts, an allegorical commentary on a poem ascribed to Orpheus. The Alexandrian school (ca. 300 BC-600 AD), centred around the Egyptian city of Alexandria, was responsible for the creation of the earliest textual commentaries, the earliest glossaries and the earliest marginal markers used to flag up particular textual features or issues (Turner, 2014, pp. 11-12). In the second century BC, Rome caught up with Greek philology and was responsible for its popularisation and transmission to later ages (Turner, 2014, p. 16). Over time, the proliferation of scribal copies of essays, commentaries, treatises and other critical works contributed to the evolution of ancient thought and language, and to, in some cases, the survival of works that are now lost. In the Late Medieval period, political and religious turmoil in Europe weakened the impetus, and the interest in classical antiquity was only to be later resumed by Italian Humanism, whose need to go back to the sources “led to the founding of philology” as we know it today (Nichols, 1990, p. 2). The advent of the (profitable) printing press in the Fifth Century reinforced this effort of textual recovery and sought to further popularise this practice. Yet, despite centuries of editorial activity and the stimulus of print, it was only in the early eighteenth century that the methods and techniques of philology were first formalised. The German Latinist Karl Lachmann (1793-1851), in fact, is known to

---

2For more information about the Derveni Papyrus, visit: [https://chs.harvard.edu/CHS/article/display/5418](https://chs.harvard.edu/CHS/article/display/5418) (Accessed: 7 December 2017).

3The *Histories* of Roman historian Paulus Orosius (fifth century AD), for instance, contains fragments of books of Cornelius Tacitus’ own *Histories*, which no longer survive.
have argued that the genealogical reconstruction of the lost original text (that is, the text as prepared by its author) could be pursued through a comparative analysis of common errors in two or more manuscript witnesses, as these provided some form of evidence of familial relation (Maas, 1958). This method, which, as was later demonstrated, had been incorrectly attributed to Lachmann by Maas and never employed by Lachmann himself (Fiesoli, 2000), assumed that in every new copy the copyist introduced new errors, that no two copyists could make the same error, and that these copies derived from a single exemplar (the archetype) (Trovato, 2014, p. 57); it therefore inherently required that the editor base the reconstruction of the text (constitutio textus) on the selection of manuscripts sharing similar errors, excluding those that seemed to deviate from the main line of transmission. Lachmann’s argument was introduced in France by Gaston Paris (1839–1903) but opposed by philologist Joseph Bédier (1864–1938), who, instead, proposed a reconstructive method based on correcting the mistakes of the best manuscript only (codex optimus), that is, the manuscript believed to be the most faithful to the original (Bédier, 1928). Bédier’s best manuscript approach echoed print’s practice of creating a fixed text in which variation was limited, not reproduced (Nichols, 1990, p. 3). Giorgio Pasquali (1885–1952) acknowledged Bédier and refuted Lachmann, advocating a method that would account for, and collate, all manuscript witnesses in the reconstruction of a text rather than a selection (Trovato, 2014, p. 72). In 1989, the debate took yet another turn with Bernard Cerquiglini’s (1947–) publication of Éloge de la variante: Histoire critique de la philologie, widely considered to be the forerunner of “New Philology” by Driscoll (2010). Cerquiglini promoted the idea of diversity by bringing the attention back to variance and away from the single-text model. In 1990, Italian philologist Gianfranco Contini (1912–1990) expressed the need for an edition to break free from the dogmatic representation of a text to, rather, reflect its mobil-

---


5For a detailed description of the history of Lachmann’s method, see Timpanaro (2005).
2.1. A short history

ity as it goes through repeated editorial assessment (Contini, 1990, p. 14). This idea of dynamic text was later recaptured by McGann (1991), who examined the issues of textual stability in the (re)production of text, and by Huitfeldt (1994), who described the tension between the multidimensional nature of texts and the one-dimensional medium that is the edition. The first overview of digital texts in the Humanities came in the year 2000 with Hockey’s volume entitled *Electronic Texts in the Humanities: Principles and Practice* (Hockey, 2000). As the title suggests, this practical introduction presented a range of tools and techniques to work with electronic texts to assist scholars with linguistic and literary analyses as well as digital editions. However, the vast potential and numerous possibilities of computational technologies described by Hockey was met with some resistance. Indeed, in 2006, textual critic Tanselle wrote:

But when the excitement leads to the idea that the computer alters the ontology of texts and makes possible new kinds of reading and analysis, it has gone too far. The computer is a tool, and tools are facilitators; they may create strong breaks with the past in the methods for doing things […] they do not change the issues that we have to cope with. (Tanselle, 2006, p. 2)

Tanselle’s provocative view of the computer in editorial practice overlooks the analytical benefits brought about by information contextualisation and aggregation, which, thanks to digital technology, can be performed at previously unthinkible scales. Indeed, while innovation in scholarly editing and in the humanities in general might be rare, the *relationability* of the data in digital editions (e.g., through Linked Open Data) increases the visibility of these efforts (Buzzoni, 2016, p. 60), thus overcoming the failure of print scholarly editions to attract and educate a wide readership (Eggert, 2016, p. 803).

As will become clearer in the upcoming chapters of this thesis, which draw on the sizeable body of literature published in the last twenty years, the innovation brought by the digital turn has not come at the expense of traditional methods of
philological enquiry but is, instead, exploring new ways of modelling editions of text.

2.2 What is an edition?

The Oxford English Dictionary defines an edition as:

One of the differing forms in which a literary work (or a collection of works) is published, either by the author himself, or by subsequent editors. (3 concr. a.)

This can be taken to imply that an edition is an author’s interpretation of a text, as Carlquist explains:

An edition is always some sort of interpretation of the [...] text and in extension also an interpretation of the social culture in which the manuscript was produced and consumed. [...] The editor tries to interpret what the text was meant to be; he or she edits a hypothetical version of the text by using a text critical and philological method. (Carlquist, 2004, p. 112)

Further, the Lewis-Short Latin dictionary traces the origin of the term ‘interpretation’ back to interpres:

I. An agent between two parties, a broker, factor, negotiator
II. An explainer, expounder, translator, interpreter
II.B. Esp., an interpreter, dragoman

The syllogistical conclusion, therefore, is that an edition is not a replacement of the original text but an intermediary and a derivative output. Derivation, however, does not exclude originality. The originality of editions and what sets them apart is the diversity with which editors choose to reproduce the text, diversity

---

based on individual knowledge, experience, expertise and vision. One effective way of capturing this concept is to think of editions as *noisy channels*.\(^8\) In his 1948 seminal work *A Mathematical Theory of Communication*, Claude Shannon, considered to be the father of information theory, introduced the *Noisy-channel coding theorem*.\(^9\) The theorem originated in Shannon’s efforts to correct noisy telephone lines, which until that point had relied on electrical signals being dispatched between transmitter and received along a network of wires. The further the signal travelled, the weaker the speech of the transmitter; boosting the speech of the transmitter would also boost the noise. Shannon’s solution to this problem converted the speech signal into a code of ones and zeros and carried it through to the receiver, which would decode the message and cleanly relay it to the recipient. The noisy channel theorem aptly, and figuratively, summarises the process of editing text (see Figure 2.1), where a transmitter (*editor*) relays the source text (*information source*) to the reader (*destination*) along with a commentary (*noise*). The noise could be described as the knowledge that leads any one editor to codify or edit the text in a given way. In this allusion the noise should not be understood as a negative contamination of the signal but as an independent element that is externally added to the message being transmitted.

**Figure 2.1:** Shannon’s *Noisy-channel coding theorem*. Diagram adapted by me from: https://upload.wikimedia.org/wikipedia/commons/f/f3/Shannon_communication_system.svg (Accessed: 1 February 2017).

---

\(^8\)This allusion has already been made in Pierazzo (2016, p. 68).

Traditionally, a print edition of a text allocates space to different components, including the reconstructed text, a critical apparatus (containing both notes and textual variants), para-textual features such as marginal notes, a preface and indices. These spaces are differently laid out depending on the publisher or the series within which a particular edition is published. Figures 2.2 and 2.3 and 2.4 below are only three examples of the many different print editions in existence.

Figure 2.2: A work by Cicero as edited in the *Oxford Classical Text* series. Numbers in the margin are used to mark sections and lines, while the bottom part of the page accommodates the critical apparatus. Source: https://archive.org/stream/CiceroDeRePublicaOxfordZetzel/Cicero%20De%20Re%20Publica%20280xford%2C%20Zetzel%29#page/n249/mode/2up (Accessed: 27 January 2018).
2.2. What is an edition?

Figure 2.3: A work by Sulpicius Severus as edited in the Corpus Scriptorum Ecclesiasticorum Latinorum edition of Latin Church fathers. Here, numbers in the margin mark sections and lines, while at the bottom of the page space is given for notes (the standalone line string of text placed between the main text and the critical apparatus) and to the critical apparatus. Source: http://archive.org/stream/corpusscriptoru12wissgoog#page/n31/mode/2up (Accessed: 27 January 2018).

Typically, the critical apparatus is found at the bottom of the page; it contains the textual variants found in other witnesses and may optionally include notes by the editor(s) (Damon, 2016). In order to fit as much information into the critical apparatus as possible, editions follow a number of conventions to abbreviate text and thus save page space; this practice, however, thwarts legibility for those who are not familiar with the jargon.

The reconstructed text is based on a transcription of the manuscript(s) under investigation. The OED definition of transcription "The action or process of transcribing or copying" (Oxford English Dictionary) is somewhat incorrect.  

Figure 2.4: A 1709 edition of a work by Greek sophist Philostratus divides the page into three main areas: one column for the Greek text, one column for the Latin translation, and a bottom section for the critical apparatus; the order of the columns is alternated between pages. Source: http://archive.org/stream/tatonphilostrato00phil#page/302/mode/2up

much as transcription is not a mere form of lifting text off a manuscript page, but a veritable editorial endeavour (Huitfeldt and Sperberg-McQueen, 2008). The importance of transcription has been extensively discussed by leading scholars in the field, all of whom maintain that the act of interpreting and commenting a text starts with, not after, transcription.\(^\text{11}\)

To transcribe a manuscript is to select, to amalgamate, to divide, to ignore, to highlight, to edit. (Robinson, 1993, p. 10)

\(^{11}\text{Including (in alphabetical order) Hans Walter Gabler, Elena Pierazzo, Peter Shillingsburg, Peter Robinson and Edward Vanhoutte.}\)
2.2. What is an edition?

Transcription may retain part or all of the information in a document, depending on the needs of the editor and on the audience the edition targets. Driscoll identified the following levels of transcription (Driscoll, 2006):

- **Diplomatic**: includes all features present in the source text; as the closest reproduction of the original, these transcriptions principally target experts.

- **Semi-diplomatic**: stays close to the original while concurrently providing explanatory information.

- **Semi-normalised**: moves away from the original by partly normalising the script.

- **Normalised**: fully normalises the text thus catering for a wider readership.

Whether diplomatic, semi-diplomatic, semi-normalised or normalised, transcribing is a time-consuming and labour-intensive activity in that each and every detail has to be accurately recorded (Fischer, 2012, p. 83). Should the transcription not be accurate, the imprecision is inescapably inherited by, and thus invalidates, any derivative work based on it. In order to produce the most accurate of transcriptions, the editor must have access to the manuscripts or at least to high quality facsimiles or photographic reproductions. Ideally, an editor is granted access to all necessary documents but institutions and libraries, as Jones informs, are not always able to support researchers in their projects when it comes to special collections, which is where manuscripts are often placed (Jones, 2004, p. 89). Even if access is granted, print editions do not provide full transcriptions of all manuscripts used for the *constitutio textus* but limit themselves to listing the variants (Fischer, 2012, p. 85). It follows that the engagement of the readers with the text is passive, insomuch as they are not able to compare the edition’s reconstructed text with the source material and are instead expected to trust the

---

12 A library’s Special Collection typically contains rare, precious, old and archival material often subject to access restrictions or supervision due to their special condition or status. For an overview of the issues surrounding access to special collections, see the most recent OCLC report (Dooley and Luce, 2010). While specific to the US and Canada, the discussion should be understood to have a wider application.
editor’s notes. But while expert readers might be able to infer the editor’s deduc-
tion process through their own knowledge of the subject matter and by consult-
ing the critical apparatus, an untrained mind is not sufficiently equipped to make these inferences. The power of the digital over print here is evident, in that in a
digital setting space is of no consequence, giving editors the freedom to dissect
and elaborate to any degree, at least in principle.

[…] the digital medium is particularly suited to providing the trans-
pparency that is fundamental for scholarly research and is imperfectly
realized by the apparatus in scholarly print editions. (Fischer, 2012,
p. 86)

2.3 What is a digital edition?

Alle Editionen seit dem Beginn dieses Jahrtausends sind digitale Edi-
tionen. (Vogeler, 2017)¹³

As previously mentioned, in the Digital Humanities the term ‘digital edition’ is
typically used to denote projects that make use of digital technologies to repro-
duce editions and transcriptions of literary texts, be those inscribed on tablets or
penned on papyrus, vellum or paper. Digital editions themselves are to be con-
sidered an asset of cultural heritage, as Tomasi explains:

Le edizioni digitali fanno parte del patrimonio culturale e vanno quindi
valorizzate al pari delle raccolte librarie, archivistiche e museali, anche
in considerazione della realizzazione di digital libraries nella forma di
aggregatori di risorse come strumento di accesso integrato al patrimo-
nio culturale [...]. (Tomasi, 2013, p. 25)¹⁴

In its broadest possible meaning, a digital edition is an edition or transcription of

---

¹³My English translation: “All editions published in this millennium are digital editions.”
¹⁴The following is my English translation of the quote: “Digital editions form part of cul-
tural heritage and should, therefore, be valued equally to library, archival and museum collec-
tions; they should also be considered for inclusion in digital libraries —intended here as resource-
aggregator—as access tools to cultural heritage [...].”
2.3. What is a digital edition?

A text that makes use of digital technologies to enhance users’ access and experience of the source material. A digital edition can either be a born-digital project or the digital rendition of an existing print edition of a text.

Vanhoutte’s definition, published in 2006, succinctly profiles digital editions:

By electronic edition, I mean 1. the immediate result or some kind of spin-off product from textual scholarship, 2. which is intended for a specific audience and designed according to project-specific purposes, 3. which represents at least one version of the text or the work, 4. which has been processed from a platform-independent and non-proprietary basis that is both stored for archival purposes and is made available for further research (Open Source Policy), 5. whose creation is documented as part of the edition, and 6. whose editorial status is explicitly articulated in the edition.

Patrick Sahle’s definition, instead, addresses the scholarly character of a digital edition:

\[ S – Scholarly: \text{ An edition must be critical, must have critical components. A pure facsimile is not an edition, a digital library is not an edition.} \]

\[ D – Digital: \text{ A digital edition cannot be converted to a printed edition without substantial loss of content or functionality. Vice versa: a retro-digitised printed edition is not a scholarly digital edition (but it may evolve into a scholarly digital edition through new content or functionalities).} \]

\[ E – Edition: \text{ An edition must represent its material (usually as transcribed/edited text) - a catalogue, an index, a descriptive database is not an edition.} \]

\[ Complete/Prototype: \text{ A SDE (Scholarly Digital Edition) is a publication of the material in question; a SDE project is not the same as a} \]
SDE, that means a SDE is more than a plan or a prototype.\footnote{Analysis presented in the About page of Sahle’s A Catalog of Digital Scholarly Editions website, available at: http://www.digitale-edition.de/vlet-about.html (Accessed: 23 October 2014).}

Sahle’s definition has led others to distinguish between digital editions and digital scholarly editions to underline the presence of a critical component in a project.\footnote{These are being collected in the Lexicon of Digital Scholarly Editing under the entry “edition (scholarly)”, at: http://uahost.uantwerpen.be/lse/index.php/lexicon/scholarly-edition/ (Accessed: 7 December 2017).} Fischer, for instance, further articulates criticism into annotation, markup, metadata and documentation, describing them as essential features of a critical text (Fischer, 2017).

A digital edition can assume different shapes, including genetic (i.e., illustrates the chronological evolution of the text), diplomatic (i.e., transcription of the text ‘as is’), facsimile (i.e., a photographic reproduction of a text with a commentary), eclectic (i.e., a varied collection of manuscripts) or critical (i.e., the editor edits the original text to make it more legible by, for instance, supplying missing words or by expanding abbreviations) (Pierazzo, 2014).

\section*{2.4 Advantages of digital editions}

Scholars in the field of digital philology typically articulate the advantages of digital editions over their print counterparts into five categories: storage capacity, relationability to other data, interoperability between different computer systems, multimediality and user interaction (Buzzoni, 2016, pp. 59-60). These can be summarily grouped into two interconnected domains, access and flexibility.

Access can be understood as both access to the primary sources and the materials that constitute the digital edition, as well as accessibility of the content. In a digital edition, the former is enhanced with the publication of digital images of the source(s) under study, which may function as a placeholder should the original item be unavailable due to conservation or loan. Where technology and rights permit, digital images can be manipulated, downloaded and reused. Access
can be further improved with the inclusion of all information or documentation needed to overcome the domain-specific accessibility issues that characterise print editions (e.g., most notably, the critical apparatus) (Eggert, 2016).

Contrary to print, where the imposed space restrictions limit the editor’s creative freedom, the flexibility of digital editions gives builders total freedom over their publication (Deegan and Sutherland, 2009, p. 6). One of the advantages of working in a digital environment is the option to gradually release content over a period of time, much like a “work-site”, helping to keep content relevant and promoting greater visibility (Eggert, 2005, 2009). For example, one can publish manuscript transcriptions and only later decide to add a linguistic layer of annotation to the text. This open-endedness of digital editions can help creators prioritise tasks and publish components independently of one another. That said, “a perpetually ‘in progress’ [electronic resource] can be quite a burden on the author or editor […]” as the continuous incorporation of new scholarship requires time and resources (Bodard, 2008).

The digital edition ages and needs continuous monitoring and maintenance whereas print has a much longer lifespan and is work-free.

(Paloff, 2011)

As previously mentioned, the larger storage capacity of the digital medium can accommodate any level of detail: for instance, instead of producing one transcription or edition of a text from a selection of witnesses, in a digital setting the editor can provide the full transcription of each witness and allow for comparison within the edition interface; an example is the New Testament Virtual Manuscript Room project, a collaborative effort to transcribe a large number of witnesses of the New Testament and which allows the user to collate any selection of manuscripts on the fly.\textsuperscript{17} Other examples are the Inscriptions of Aphrodisias project, which provides users with different textual layers for each document or object,\textsuperscript{18} and the ABaC:us project, whose texts are accompanied by linguistic

\textsuperscript{17}Available at: http://ntvmr.uni-muenster.de/collation (Accessed: 27 January 2018).
\textsuperscript{18}See, for example, the three different textual layers (Edition; Diplomatic; Epidoc XML)
A good example of a large storage use is the *Chopin’s First Editions Online* project (see Figure 2.5), which gives users the option to compare a large number of images: depending on resolution, dimensions and format, digital images can be upwards of 10 megabytes in size, thus potentially requiring gigabytes worth of storage space.

![Figure 2.5: The Chopin’s First Editions Online project (http://www.chopinonline.ac.uk/cfeo/) allows users to choose editions to compare; the images can be enlarged for a detailed study of the similarities and differences. This screenshot shows a comparison of two editions of *Rondo pour le Pianoforte, Op.1*.](image)

Other possibilities granted by the flexibility of digital editions are the integration of contextual information, timelines, videos and any information the editor provided for the inscription here: http://insaph.kcl.ac.uk/iaph2007/iAph018021.html (Accessed: 27 January 2018).

19 Available at: https://acdh.oeaw.ac.at/abacus/ (Accessed: 27 January 2018).

20 For a comparison and estimation of image formats, see for example: http://www.digitalmemoriesonline.net/scan/output/jpeg_vs_tiff.htm (Accessed: 27 January 2018).
judges relevant for a comprehensive understanding of the source text (e.g. the Livingstone Online project);\textsuperscript{21} the use of animation for a dynamic exploration of the source materials (e.g., the Digital Dead Sea Scrolls project, as seen in Figure 2.6);\textsuperscript{22} the use of Linked Open Data to build connections between the digital edition and other online resources (e.g., the J. C. Senckenberg (1707-1772) Diaries project’s links to the Deutsche Nationalbibliothek Katalog and relevant Wikipedia pages);\textsuperscript{23} the possibility of citing a digital edition as a whole or to assign different identifiers to individual items or components (e.g., the Getty Pietro Mellini Inventory in Verse 1681 project);\textsuperscript{24} the possibility of licensing the components of the digital edition differently (e.g., XML transcription available for download but not images, as is the case of the Corpus of the Inscriptions of Campá);\textsuperscript{25} the possibility of opening up a project to community engagement (e.g., the Social Edition of the Devonshire MS (BL Add. MS 17492));\textsuperscript{26} the possibility of virtually re-uniting dispersed manuscript fragments (e.g., the Codex Sinaiticus project);\textsuperscript{27} of accommodating content contributed by users;\textsuperscript{28} of reducing the complexity foisted by the critical apparatus upon the small readership of traditional scholarly editions through textual expansion, hyperlinking functionality as well as visualisation (Buzzoni, 2016, pp. 64-67); and, finally, the possibility, at least in principle, of modifying content at any time and integrating user feedback (Eggert, 2009, p. 64).\textsuperscript{29} Open-endedness also lends itself well to collaborative editing and teach-

\textsuperscript{21}Available at: www.livingstoneonline.org/ (Accessed: 27 January 2018).
\textsuperscript{22}Available at: http://dss.collections.imj.org.il (Accessed: 27 January 2018).
\textsuperscript{23}An example can be viewed here: http://sammlungen.ub.uni-frankfurt.de/urn/urn:nbn:de:hebis:30-83754 (Accessed: 27 January 2018).
\textsuperscript{24}An example can be viewed here by clicking on 'Cite': http://www.getty.edu/research/mellini/manuscript/5-recto (Accessed: 27 January 2018).
\textsuperscript{26}Available at: https://en.wikibooks.org/wiki/The_Devonshire_Manuscript (Accessed: 27 January 2018).
\textsuperscript{28}An example of this is Papyri.info. Or, more notably, the crowd-sourcing initiative Transcribe Bentham, available at: http://blogs.ucl.ac.uk/transcribe-bentham/ (Accessed: 2 February 2017).
\textsuperscript{29}For example, the Henry III Fine Rolls Project (Available at: http://www finerollshenry3.
ing. For example, Brown’s Decameron Web project team firmly believe that their
digital edition enables teachers and learners to browse the texts in collaborative,
‘playful’ ways, thus adding a new dimension to the more traditional approach
to manuscripts. Another project by Brown University, the Women Writers Project,
invites teachers and the general public to submit assignment and lesson ideas in-
volving the project’s digital resources and texts with a view to creating a more
engaging learning experience. Finally, the open-endedness of a digital edition
can also be conducive to the creation of more editions (Ore, 2009, p. 118).

![Figure 2.6: The Great Isaiah Scroll of the Digital Dead Sea Scrolls project can be (un)rolled and read in a dynamic fashion using the text-image link and macro-view features. Source: http://dss.collections.imj.org.il/isaiah#2: 4 (Accessed: 27 January 2018).](image)

2.4.1 Infrastructure for digital editions

For all its merits, the plasticity of digital editions imposes technological and modelling choices capable of accommodating the volatility of data while adhering to established methods of practice (Buzzetti and Rehbein, 2008; Ciula and Eide, 2014; Sahle, 2012; Pierazzo, 2013). The choice of tool(s) to be used in the creation of a digital edition depends as much on project scope as it does on the usability, reliability and mutability of a given technology. A survey conducted by Gibbs and Owens reveals that, generally, humanities scholars working in the field of Digital Humanities judge and select digital tools by their ease of use, their documentation and overall presentation (Gibbs and Owens, 2012). In fact, the authors found that humanists’ scepticism towards digital tools is not dictated by a reluctance to learn how to use those tools but, on the contrary, by the time necessary to understand what the tools do and if they help to address their research questions:

There are plenty of interested, curious, and technically capable humanities researchers that have little time and patience for trial and error with new methodologies when they are uncertain of their value. However, they remain receptive to the possibilities offered by the tools. (Gibbs and Owens, 2012)

Infrastructural decisions for digital editions are particularly challenging, inasmuch as creators must assess a given tool’s adequacy for a particular task or component of the project in relation to all other tools and components. Although the proliferation of technology gives one more choice, the integration of different tools often requires adjustments or custom intervention to bridge technologies and thus streamline work processes. To narrow this choice and reduce the infrastructural complexity of digital editions, some professionals have developed ad hoc platforms and tool-sets. One such example is Omeka, a ‘free, flexible, and open source web-publishing platform for the display of library, museum, archives, and scholarly collections and exhibitions’, as stated on the offi-
cial website.\(^{31}\) Another example is TextGrid, a digital environment built to support researchers in the creation of digital editions of large corpora or individual works;\(^{32}\) through TextGrid, users can share and store files, collaboratively annotate and link text to digital images. Other examples include the increasingly popular *Edition Visualization Technology* (EVT), a bundle developed by college students and designed to manage the file transformation and visualisation process of both images and XML-encoded text automatically (see Figure 2.7);\(^{33}\) King’s College London’s *Kiln* platform, an open source solution for the publication of XML source content;\(^{34}\) the *Generic-de-web-app* application developed by the Austrian Centre for Digital Humanities, which promotes maintenance, interoperability and rich customisation;\(^{35}\) and LombardPress, an extensible framework with Linked Open Data and IIIF (*International Image Interoperability Framework*) integration geared towards the creation of scholastic commentaries and texts.\(^{36}\)

Many of these technologies rely on version control systems, such as *GitHub*,\(^{37}\) *GitLab*,\(^{38}\) *Sourceforge* or *Apache Subversion*,\(^{39}\) to track, and optionally share, the progress of a project.\(^{40}\)

\(^{31}\) *Omeka* is a project of the Roy Rosenzweig Center for History and New Media, George Mason University. Available at: http://omeka.org/ (Accessed: 11 January 2012).


\(^{33}\) Available at: https://sourceforge.net/projects/evt-project/ (Accessed: 2 February 2017).


\(^{35}\) As stated on the project website here: https://www.oeaw.ac.at/acdh/tools/generic-de-web-app/ (Accessed: 27 January 2018).


\(^{37}\) Available at: https://github.com/ (Accessed: 15 September 2014). For an example on how GitHub is being used in the Humanities, particularly book publishing, see http://go-to-hellman.blogspot.de/2015/01/why-github-is-important-for-book.html (Accessed: 1 February 2015).

\(^{38}\) Available at: https://about.gitlab.com/ (Accessed: 27 January 2018).


\(^{40}\) For a more detailed overview of the adoption of version control systems in Humanities scholarship, see Brown and Simpson (2014).
2.4. Advantages of digital editions

Figure 2.7: The Digital Vercelli Book project makes use of the Edition Visualization Technology (EVT) to publish XML-encoded text and manuscript images in the web in a functional reading environment.

2.4.1.1 Extensible Markup Language (XML) and the Text Encoding Initiative (TEI)

As Chapter 3 of this thesis shows, many digital editions employ XML (Extensible Markup Language), a descriptive markup language widely adopted in the Digital Humanities to produce machine-actionable and searchable text (McDonough, 2009). While part of the appeal of XML is complete freedom over the choice of human-readable and intuitive tags to describe text (e.g., the name ‘Jane’ could be tagged as <name>Jane</name> or <firstName>Jane</firstName> or <femaleName>Jane</femaleName>), in 1987 a community of scholars known as the Text Encoding Initiative (TEI) introduced guidelines to enforce some degree of mark-up consistency with a view to enhancing the interoperability of XML-encoded documents (Vanhoutte, 2004) (Barney, 2012). The TEI website states:

The Text Encoding Initiative (TEI) is a consortium which collectively develops and maintains a standard for the representation of texts in digital form. Its chief deliverable is a set of Guidelines which specify encoding methods for machine-readable texts, chiefly in the human-
Chapter 2: Digital editions

ities, social sciences and linguistics. Since 1994, the TEI Guidelines have been widely used by libraries, museums, publishers, and individual scholars to present texts for online research, teaching, and preservation.\footnote{Available at: \url{http://www.tei-c.org/index.xml} (Accessed: 30 October 2014).}

Within this standard, Special Interest Groups (SIGs) were set up to discuss and devise narrower sets of TEI markup suitable to particular typologies of text (e.g., correspondence or music).\footnote{The complete list of SIGs is available at: \url{http://www.tei-c.org/Activities/SIG/} (Accessed: 30 October 2014).} Similarly, a number of subsets were created to deal with distinct project purposes: TEITite, for example, is a TEI customisation often encountered in a library setting and used in digitisation projects, which outsource work (Barney, 2012, p. 38); EpiDoc is another subset originally conceived to digitally reproduce epigraphic and papyrological documents but its tag-set also lends itself well to ancient manuscripts.\footnote{For more information about EpiDoc, visit: \url{http://sourceforge.net/p/epidoc/wiki/Home/} (Accessed: 15 September 2014).}

2.4.2 Challenges of digital editions

Although TEI is widely adopted, Chapter 3 also shows that many projects do not employ this standard. Reasons given for this include the inadequacy of XML-TEI for particular use-cases or the refusal to learn a standard (Barney, 2012, p. 39), which is not as interoperable as it claims to be Schmidt (2014). For instance, two XML-TEI-encoded documents from two different projects might use tags in identical ways but do not necessarily mark-up text in the same level of detail; to make these two documents consistent, the editor must identify the differences and adapt the markup accordingly.

There has never been a single standard convention for the transcription of manuscripts texts, and it is not likely there will ever be one, given the great variety of textual complications that manuscripts— from all times and places— can present. (Meulen and Tanselle, 1999,
Providing a number of different formats for a single dataset increases the chances of reuse and interoperability (Schmidt, 2014, p. 14).

As well as technical difficulties, digital editions pose durability, accessibility, sustainability and financial challenges (Cohen and Rosenzweig, 2006).

Critical editions require considerable effort on many fronts, starting from the development of tools to support the many operations that are part of producing such an edition, such as transcription, collation, stemma generation, and web publishing, among others. (Pierazzo, 2014, p. 13)

As Warwick notes, the range of skills necessary to create a digital edition are rarely found in a single individual:

[the textual editor] is very likely to be a remarkable polymath, adept in the traditional skills of textual criticism and transcription as well as the new technologies involved with electronic publication, or more likely, she will be more than one person. (Warwick, 2001, p. 56)

Still, an interdisciplinary and multi-skilled team is not without difficulties either:

L’atto della creazione di un’edizione elettronica è processo che si configura come luogo d’incontro e confronto dialogico e, come spesso capita, anche di conflitto e incomprensione tra diverse culture e pratiche, in primis tra il pensiero narrativo e quello scientifico [...]. (Meschini, 2013, p. 29)

Another challenge of digital editions is inclusivity. While scholars argue that digital editions should not and cannot meet the demands of every user (Fischer, 2017, p. S281) (Zundert, 2012, p. 172), some user responses given to the survey described in Chapter 4 of this thesis inform that no digital edition meets their
needs. And yet, to cater for everyone, creators might be pushed toward generalisation, thus moving away from the specificity which might be useful to some (Zundert, 2012, p. 172) (Meschini, 2013, p. 38). Bridging the gap between creators and user might prove an impossible challenge to conquer, but, as Chapter 4 later proposes, by better defining their stance with respect to a given work (i.e., the peculiarities of their project, their research question(s) or perspective, as well as their intended audience), creators of digital editions have the opportunity to help their readership better appreciate and use the results of their efforts.

Versionality and citability also pose a significant challenge. Indeed, the progressive nature of digital editions calls for systems capable of not only keeping track of versioned editing but of project documentation and website updates as well (Fischer, 2017, p. S286). Modern research data repositories such as Zenodo\textsuperscript{45}, Figshare\textsuperscript{46} and GitHub facilitate the addressability of progressive publication but their slow uptake in digital editing (see Chapter 3) might be symptomatic of incompatible data and infrastructure models.

Perhaps the thorniest challenge of all is project maintenance and sustainability. As Eggert recently noted, “The need for the edition to last 100 years […] simply does not obtain in the digital medium” (Eggert, 2016, p. 804), and despite the establishment of dedicated centres to help fight technological obsolescence, and help maintain and support the long term sustainability of research output,\textsuperscript{47} the durability and longevity of digital editions remain an open issue.

I would argue that these very challenges contribute to the attraction of working in this medium […]. Electronic editing can be daunting financially, technically, institutionally, and theoretically— but it is also a field of expansiveness and tremendous possibility. (Price, 2008)

One might ask: so if digital editions are not sustainable, why create them and what is their real value? This is a valid question but one that can be asked about

\textsuperscript{45}Available at: \url{https://zenodo.org/} (Accessed: 14 May 2018).
\textsuperscript{46}Available at: \url{https://figshare.com/} (Accessed: 14 May 2018).
\textsuperscript{47}For example, the Humanities Data Centre at the University of Göttingen in Germany.
any (digital) resource being produced today. I would argue that the value of a resource is not only defined by its longevity but also by the opportunities it offers to both creators and users during its lifetime. As will become clearer later in this thesis, to create a digital edition is to frame humanistic enquiry in a digital context. In doing so, creators reassess knowledge in a different light, giving anybody with web access the opportunity to join this process of rediscovery.

2.4.3 The status quo

This tremendous possibility stimulates experimentation and has summoned a vibrant digital editing community of practice engaged in the creation of resources (e.g., Lexicon of Scholarly Editing), standards (e.g., Text Encoding Initiative), journals (e.g., RIDE Review Journal of Digital Editions and Resources, Scholarly Editing, Digital Scholarship in the Humanities (formerly Literary and Linguistic Computing), evaluation criteria (e.g., IDE Criteria for Reviewing Digital Editions), courses and training programmes (e.g., Digital Scholarly Editions Initial Training Network, DARIAH Digital Scholarly Editions: Manuscripts, Texts, and TEI Encoding) and catalogues (e.g., Patrick Sahle’s Catalog of Digital Scholarly Editions and my own, as described in Chapter 3).

Indeed, the influence exerted by internet technologies upon those working in the field of digital philology is inducing a metamorphosis of digital editions, aptly captured by Jeffrey Witt in these terms:

The publication of DSEs [Digital Scholarly Editions] is not best iden-
tified with a publication of a book or website […] rather the public-
lication of a DSE should coincide with the publication of a dataset,
structured and serialized and made accessible according to the best
practices of the field and Linked Open Data. (Witt, 2016)

Yet, despite the sizeable literature on digital editions published over the past
forty years, it remains unclear how these are being created and to what extent
creators are meeting user needs. Chapters 3 and 4 address these questions and
represent the largest, most comprehensive studies yet.
Chapter 3

A Catalogue of Digital Editions

Since the earliest days of hypertext (1980s), textual scholars have produced and discussed digital editions of manuscript culture, in order to investigate how digital technologies can provide another means to present and enable the interpretative study of text. However, this work has generally been done by looking at particular case studies or examples of digital editions, so there is no overarching understanding of how digital technologies have been employed across the full range of textual interpretations and, consequently, of how the field is evolving and in which directions it may be heading.

The application of digital technology to the study and production of texts has positively affected the way researchers communicate their research to their audiences and peers, but has also given way to a wide range of reproduction methods to the point where no two digital editions are the same unless they are produced by the same individual or team. While the traditional methods of philological enquiry still hold true and are widely used among digital editors, the choices made to reproduce texts in a particular way are more difficult to pin down. To better understand how the theory and practice of digital editing come together, I have built a Catalogue of Digital Editions, which provides a means of answering in the form of a quantitative and qualitative overview of digital editions the following research questions: what makes a good digital edition? What features do digital editions share and how do they differ? In light of the different tech-
nological expressions one can observe across digital editions, the Catalogue can be used to inform the definition of the minimum requirements a digital edition should fulfil, and to help build future projects.

This chapter begins by outlining the motivation for building the Catalogue of Digital Editions and relevant research. It presents the method used to gather the data and describes the development of a web application to increase the usability of the data-set. Next, it discusses the limitations and the sustainability challenges faced by the Catalogue as a digital resource, and, finally, details the achievements, results and future steps of the research.

This chapter is based on a continuously growing published digital resource. The snapshot described here is based on data collected as of May 2017.

3.1 Why build a Catalogue of Digital Editions?

The idea to build a Catalogue of Digital Editions took shape in September 2012,\(^1\) at the *Quarto Incontro di Filologia Digitale* in Verona, Italy,\(^2\) when Professor Caroline Macé expressed the need for a database of digital editions to facilitate the discovery, linking and advertising of digital editions of text, which may otherwise go unnoticed. That database would give an insight into past and present efforts, providing the means to identify and view patterns of practice and areas for improvement, and preserving the memory of discontinued projects. The database could inform the design of new digital editions and serve as a tool for their popularisation.

The concept of such a database, however, was not new. In 2012, lists of digital editions already existed. The first, large database to have been produced is Patrick Sahle’s *Catalog of Digital Scholarly Editions*,\(^3\) which currently re-

\(^1\)An earlier discussion of the *Catalogue of Digital Editions* is provided by Franzini et al. (2016).
\(^2\)For more information about the event, see https://folk.uib.no/hnooh/filologiadigitale/index.html (Accessed: 20 April 2018).
\(^3\)As the name suggests, the Catalog lists only scholarly editions, that is, projects carrying a strong critical component. Personal correspondence (14/06/2012) with Sahle revealed that the
3.2 Method

In the first phase of the project, the data was collected and catalogued in two ways: through direct acquisition and consultation of project websites, and from a number of project creators by way of a short questionnaire. This section begins with a description of the set-up of the Catalogue and subsequently turns to the


Monella’s catalogue is available at https://docs.google.com/document/d/1rmCkvtVJmLcJrJSUOx90dSEcgs7M003JdLdEb7, section 2.2 (Accessed: 12 February 2013);
Pusceddu’s at http://www.digitalvariants.org/e-philology (Accessed: 12 February 2013);
qualitative and quantitative research methods used.

### 3.2.1 Building the Catalogue of Digital Editions

The Catalogue of Digital Editions consists of a single spreadsheet. While primarily populated by me, the file was always intended to accommodate external contributions as well. The first version was published in May 2013 and was built with Google Sheets (Franzini et al., 2016). This technology worked well for a small amount of data (i.e. approximately 100 digital editions with 20 features or columns each) but the growing number of digital editions and the increasing complexity of the cataloguing significantly affected the stability and reliability of Google Sheets. Not only did loading times become slower but the absence of administrative control over the spreadsheet could not counter the accidental data deletion or overwriting caused by the occasional user contribution. For these reasons, in 2015 the Catalogue was moved to GitHub, an open source development platform, which, among other benefits, provides advanced versioning functionality and complete administrative control over collaborative projects.\(^5\)

The GitHub repository of the Catalogue of Digital Editions is available at [https://github.com/gfranzini/digEds_cat](https://github.com/gfranzini/digEds_cat) and has been given a Digital Object Identifier (DOI) for citation purposes. Figure 3.1 shows the repository’s homepage.

The use of GitHub requires the creation of a user account. GitHub offers both basic (free) and premium (paid) accounts depending on the type of services needed. The Catalogue was set-up using a free GitHub account. Users of the Catalogue who wish to contribute data must also create a GitHub account, otherwise the resource can be browsed freely without registration.

A GitHub project is known as a repository and every GitHub repository created with a free account comes with eight default information layers (Figure 3.2): Code displays the code or the files that make up the repository; Issues provides a means of reporting bugs or assigning tasks; Pull requests regulates collaborative

---

\(^5\)GitHub is available at: [https://github.com/](https://github.com/) (Accessed: 10 April 2017).
3.2. Method

Figure 3.1: GitHub repository of the Catalogue of Digital Editions (September 2017 snapshot).

work; Projects lists the projects within the repository; Wiki provides a space for documentation; Pulse and Graphs monitor activity in the repository; and Settings allow users to customise the repository in various ways.

Figure 3.2: Repositories created with free GitHub accounts come with seven information layers (September 2017 snapshot).

What follows is a description of the Catalogue’s Code, Issues and Wiki layers
only as all others are not relevant to this discussion.

Figure 3.3 illustrates the Code layer of the repository. It lists all files that make up the Catalogue and when these were last modified. Like many cloud services, a GitHub repository is hosted on a remote server but can be managed locally as a traditional file folder. Any changes made locally need to be synchronised with the server in order to be viewed by users; to work with GitHub, one needs to either install a designated Graphical User Interfaces (GUIs) or invoke Git commands using the command line. To contribute to a repository as an external user the repository must be copied or cloned; every modification or addition made to the copy can be then synchronised with the master repository through a moderation process. The core files of the Catalogue’s GitHub repository are two .csv files: `digEds_cat.csv` collects digital editions and `institutions_places_enriched.csv` collects the geographical coordinates of the producing institutions recorded in `digEds_cat.csv` (i.e. institutions involved in the creation of a digital edition, not the funding agencies - not yet, at least).\(^6\)

The choice to use the Comma Separated Value (.csv) file format was dictated by its versatility and compatibility across different calculus software. It can be ingested by Microsoft Office or Libre Office and can be exported to many other file formats. The other files within the repository include a `README.md` file, used to introduce the project and a `CONTRIBUTING.md` file, which contains instructions on how to contribute digital editions to the Catalogue.

Users can freely open and download files, but can also explore each and every modification that was made to a particular file. If, for example, a user were to click on the ‘Corrected entry #218’ commit message that appears next to the `digEds_cat.csv` file in Figure 3.3, the resulting page would display the correction, as shown in Figure 3.4. Here, the user can clearly see a change in line 218 of the file. GitHub uses red to highlight the line as it was before the change, and green to highlight the line after the change. Stronger hues of red and green are used to point the user to the exact location of the correction within that line, in

\(^6\)The reason for recording coordinates will become clear in section 3.3.
3.2. Method

Figure 3.3: The Code layer of the Catalogue’s GitHub repository, containing all the files that make up the Catalogue (April 2017 snapshot).

In this case the year, which changed from 2014 to 2017. This is what is known as a ‘versioning system’, in that it shows the version history of a file. Every correction is assigned an ID so as to be able to trace every change ever made to a file during its GitHub life.

Figure 3.4: GitHub’s versioning system allows users to track all changes made to a file (April 2017 snapshot).

The Issues layer or page of the repository is used to report bugs, assign tasks and list digital editions to be added to the Catalogue. Figure 3.5 is an extract of the Catalogue’s Issues page. Issues are automatically given an ID (beginning with #) and include information such as date of creation, issuer, assignee, as well as a
label to sort and describe the nature of the issue. Issues provide a space for comment and discussion and can be attached to a project milestone (e.g., a future release or version of the project). Users can comment on a particular issue. If they do, a speech bubble icon will appear next to the issue in question.

![Image](https://example.com/image.png)

**Figure 3.5:** The Issues page of the Catalogue lists issues and tasks (September 2017 snapshot).

Finally, the Wiki describes the purpose and content of the Catalogue, how the data is structured, and provides instructions on how to contribute digital editions. Figure 3.6 below shows the landing page of the Wiki. A right sidebar functions as a menu to explore the different pages within the Wiki and a short note placed beneath the Wiki header (in the case of 3.6, under ‘Home’) dates the last edit made.

![Image](https://example.com/image.png)

Let us now turn to the data itself, which is stored in the `digEds_cat.csv` file. When opening the file in GitHub, the user is presented with the page shown in Figure 3.7. GitHub automatically renders .csv files as searchable, scrollable tables, helping users identify information more efficiently.

![Image](https://example.com/image.png)

As previously mentioned, since the Catalogue’s inception in 2012, the number of cataloguing features has grown in response to my own needs and of those who use the Catalogue in their work. Today (September 2017), the Catalogue counts 49 features or columns of information: Historical Periods; Time/Century; Edition name; URL; Scholarly; Digital; Edition; Language; Writing Support; Begin date; End date; Manager or Editor; Institution(s); Audience; Philological statement; Account of textual variance; Value of witnesses; XML-TEI transcription; XML-TEI available to download; Images; Zoom images; Image manipulation (brightness,
3.2. Method

Figure 3.6: The landing page of the Catalogue’s GitHub Wiki (September 2017 snapshot).

Figure 3.7: The digEds_cat.csv file as displayed by GitHub. The file can be searched using the search bar provided and a scrolling option allows users to move across the spreadsheet both vertically and horizontally (April 2017 snapshot).

rotation, etc.); Text-Image Linking; Source text translation; Website language; Glossary; Indices; String-matching search; Advanced search; Creative Commons License; Open source/Open access; Linked Open Data; API; Crowdsourcing;
Feedback; Technological statement; Links to external resources; OCR or keyed; Mobile-friendly/application; Print-friendly view; Print facsimile (complementary output); Repository of source material(s); Place of origin of source material(s); Sponsor/Funding body; Budget (rough); Infrastructure; Current availability; RIDE review; Sahle Catalog. These features were established based on a comparative study of a number of digital editions and as the project progresses new features might be added (e.g., a field for citation information).

The digital editions constituting the Catalogue come from numerous sources and their selection follows basic criteria: the electronic texts can be ongoing or complete projects,\textsuperscript{7} born-digital editions as well as electronic reproductions of print volumes.\textsuperscript{8} Sources include existing lists of projects, such as Projects using the TEI and the previously mentioned catalogues,\textsuperscript{9} publications (articles, reviews and books), mailing lists, social media, word of mouth, serendipitous web browsing, as well as (occasional) user contributions.

\subsection*{3.2.2 Content analysis: data collection and evaluation}

The creation of the Catalogue of Digital Editions followed a content analysis procedure, that is, a qualitative analysis of digital editions to deepen my own understanding of the field in general, and of how its theories and methods are instantiated in the projects currently available on the web (Erlingsson and Brysiewicz, 2017; Hsieh and Shannon, 2005; Krippendorff, 2004). This procedure was, and still is, complex.

In 2012, as the Catalogue was taking shape, the data categories and collection process were continuously revised as I became more deeply acquainted with the field of digital scholarly editing and the considerable diversity between digital editions began to emerge. Initially, to help establish a structure, two data collection experiments were run in parallel: a questionnaire, meant to elicit informa-

\footnotesize{\textsuperscript{7}Still available on the web.} \\
\footnotesize{\textsuperscript{8}A \textit{born-digital} edition of a text describes a project designed for digital publication.} \\
\footnotesize{\textsuperscript{9}Available at: \url{http://www.tei-c.org/Activities/Projects} (Accessed: 2 March 2017).}
tion from the creators of digital editions directly; and my own analysis of digital editions.

With regard to the questionnaire, a number of digital editions were collected and their creators contacted between August 2012 and March 2013. The short questionnaire was email-based, and sought to learn more about what the goals and achievements of the project were, what type of user enquiries or requests had been received, if any user statistics had been gathered, the project’s target audience, budget and team size,\(^{10}\) what lessons were learnt from the project, as well as sustainability issues in making the resource available in the longer term. Of the seventy-eight people contacted,\(^ {11}\) thirty-seven replied (some in full, some fragmentarily). Of the remaining forty-one, six emails bounced back due to expired email addresses and thirty-five did not reply.\(^ {12}\) With the exception of budget, which was only addressed by two respondents,\(^ {13}\) all questions were answered. Such a small number of replies could not be considered statistically relevant or representative of the digital scholarly editing community practice but served nevertheless as the first building block towards the establishment of the Catalogue’s structure.

Parallel to the questionnaire, I studied a selection of digital editions, tabulating common features or properties. This investigation engendered the development of the Catalogue of Digital Editions and, to this day, remains the primary mode of data collection. The addition of a digital edition to the Catalogue can take between 15-25 minutes depending on the exhaustiveness of the project website, and follows the guidelines available in the GitHub Wiki mentioned above. Figure

---

\(^{10}\)Given the potentially sensitive nature of the question, interviewees were given the option to not respond.

\(^{11}\)The number of emails sent out to project managers is smaller than that of digital editions as, on occasion, project managers are either co-investigators or work on multiple projects at the same time (in these instances, the questionnaire email addressed all relevant projects).

\(^{12}\)These replies were never published for anonymity reasons (information was so project-specific that merely removing the respondent’s name would not guarantee complete anonymity).

\(^{13}\)When asked about funding, one of the investigators of the *Orlando Furioso Hypertext Project*, at [http://stel.ub.edu/orlando](http://stel.ub.edu/orlando), replied (in Italian): "Su questa questione, capisce che non voglia dare forse i dati concreti" [My translation: With regards to this matter, I am sure you can appreciate why I prefer not to disclose concrete numbers] (10th March 2013).
3.8 shows an example of a digital edition in the Catalogue: in accordance with the information published on the project’s website, *Claudii Claudiani Carmina Latina* was catalogued as a project dealing with antique texts in Latin and as a *scholarly* digital edition owing to its strong critical character (the value 1 is used in lieu of “yes”; 0 is used for “no”). The feature list of the project (not visible in Figure 8 for space reasons) extends beyond the screen boundaries and acts as a vocabulary of descriptive properties. If project websites do not provide the information needed to populate these properties, the words ‘not provided’ are used.

![GitHub .csv view](image)

**Figure 3.8:** Example digital edition in the Catalogue as displayed by the GitHub .csv view.

Both approaches, questionnaire and my own analysis, were *qualitative* methods of data collection: the former, while imposing a set of questions, was designed to directly glean from creators individual experiences and motivations for starting the project; the latter is also qualitative insofar as it is based on a careful project evaluation (Gorman and Clayton, 2005, p. 3, 7). Moreover, the data collection process and content analysis of the Catalogue is *summative* because “it starts with identifying and quantifying certain words or content in text with the

---

14 As previously mentioned, the Catalogue of Digital Editions makes a distinction between scholarly and non-scholarly digital editions, and replicas of existing print volumes. If the project in Figure 3.8 did not contain any critical assessment of the text(s) at hand, the ‘Scholarly’ cell would have 0 as its value.

15 The full list of features is: Historical Period; Time/Century; Edition name; URL; Scholarly; Digital; Edition; Language; Writing support; Begin date; End date; Manager or Editor; Institution(s); Audience; Philological statement; Account of textual variance; Value of witnesses; XML-TEI Transcription; XML(-TEI) available to download; Images; Zoom images; Image manipulation (brightness, rotation, etc.); Text-Image Linking; Source Text Translation; Website language; Glossary; Indices; String matching; Advanced search; Creative Commons License; Open source/Open access; Linked Open Data; API; Crowdsourcing; Feedback; Technological statement; Links to ext. resources; OCR or keyed?; Mobile friendly/application; Print-friendly view; Print facsimile (complementary output); Repository of source material(s); Place of origin of source material(s); Sponsor/Funding body; Budget (rough); Infrastructure; Current availability; RIDE review; Sahle Catalog.
3.2. Method

purpose of understanding the contextual use of the words or content.” (Hsieh and Shannon, 2005, p. 1284). In the Catalogue’s case, the “text” is the digital edition, the “words” or “content” are the features or properties. The frequency of a particular feature across the entire set of digital editions, in turn, facilitates the detection of diverse patterns of practice, thus stimulating quantitative research into current trends and future directions. Notwithstanding, as results of subjective evaluation, summative content analysis and qualitative research in general question the reliability and validity of the data collected (Gorman and Clayton, 2005, pp. 24-25). With regard to reliability, the Catalogue only records information that is explicitly mentioned in digital edition websites. With regard to validity, the data in the Catalogue is 100% valid at the time a digital edition entry is created, but owing to the conventional nature of digital editions as works-in-progress, that initial validity ought to be regularly revised to account for any changes and additions. Therefore, in the absence of an automated means of retrieval of project updates, the Catalogue must rely on the manual efforts of its editors and users to ensure all data is current and accurate. So, although primarily a qualitative enterprise, the Catalogue also gives way to quantitative research, in that the data can be numerically and statistically surveyed. In conclusion, the research method underpinning the Catalogue is intentionally and necessarily mixed in order to comprehensively understand the heterogeneity of digital editions and to spawn new qualitative and quantitative research.

3.2.2.1 User contributions

As previously mentioned, the set-up of the Catalogue lends itself well to user contributions. Users can choose one of three ways of contributing digital editions: 1) if familiar with GitHub, they can make a copy of the Catalogue repository, edit the .csv file(s), and submit their changes for moderation and inclusion in the master repository; 2) if they are not confident users of GitHub, they can post the URL of the digital edition as a GitHub issue for my attention; 3) altern-
atively, they can fill-in a Google Form. These three options were designed to encourage and accept contributions from a wide range of users, some of whom may not wish to register a GitHub account in order to submit data. While the Google Form is the most popular form of contribution, the Catalogue has received submissions from all three systems. Since 2012, approximately fifteen individuals have contributed digital editions to the Catalogue.

3.3 The Web Application

In the summer of 2016, developers Peter Andorfer and Ksenia Zaytseva of the Austrian Centre for Digital Humanities (ACDH) in Vienna contacted me with a web application they had been developing designed to provide a searchable version of the Catalogue of Digital Editions. Their motivation for creating this application was to test a scalable web technology that the ACDH was developing, also in GitHub, to host a wide range of data-sets and projects. The Creative Commons License covering the Catalogue made it possible for the ACDH to ingest the data and evaluate the performance of this web technology. The successful test of the application kindled a collaboration between myself and the ACDH. The Catalogue of Digital Editions Web Application is available at:

https://dig-ed-cat.acdh.oeaw.ac.at/

3.3.1 Architecture and data visualisations

The Catalogue of Digital Editions Web Application is made-up of two independent and interacting components, each stored in its own GitHub repository: 1) the data, and 2) the web application. The two components are connected via a custom script that regularly retrieves and delivers the latest updates from the data repository to the web application. The project is maintained by a team of three

---

16 Available at: https://goo.gl/forms/4Ya3jwRCBl0V6xx2 (Accessed: 19 April 2017).
17 For a list of resources and tools (co-)developed by the Austrian Centre for Digital Humanities (ACDH) see http://www.oeaw.ac.at/acdh/en/resources (Accessed: 8 April 2017).
18 An example of another ACDH resource based on the same web technology is Tote Tiroler, available at: https://totetiroler.acdh.oeaw.ac.at/ (Accessed: 8 April 2017).
people: I curate the data and the web content, while Peter Andorfer and Ksenia Zaytseva manage the development of the web application.

The web application is a Django Web Framework, an open source solution written in Python.\(^{20}\) The choice of framework was made by the ACDH within the scope of its Digital Humanities programme and strategy. The web application is released under a Creative Commons Attribution Share-Alike License (BY-SA) and its GitHub repository is available at https://github.com/acdh-oeaw/dig_ed_cat

The collaboration began with the design of a logo to brand the project (see Figure 3.9). The logo is shaped as a check list of three items, Catalogue, Digital and Editions. The grey and black colour coding of the text separates the definition of the project, Catalogue, from its matter, Digital Editions. The colour scheme of the check boxes alludes to the logo of the UCL Centre for Digital Humanities, my home institution.

![Logo of the Catalogue of Digital Editions (my design).](image)

**Figure 3.9:** Logo of the Catalogue of Digital Editions (my design).

### 3.3.1.1 Architecture

The website of the web application makes use of the Bootstrap framework (HTML, CSS and JavaScript).\(^{21}\) Bootstrap was chosen because it is open source, popular in web development, well supported and documented, responsive and mobile-friendly.\(^{22}\)

---


\(^{22}\)For extra assurance, this web application was tested against, and passed, the Google Search Console Mobile-Friendly Test, available at: [https://search.google.com/search-console/mobile-friendly](https://search.google.com/search-console/mobile-friendly) (Accessed: 17 May 2017).
Figure 3.10: Homepage of the web application of the Catalogue of Digital Editions (April 2017 snapshot).

Within the web application, a top navigation menu is split into two areas. The left-hand side of the menu lists pages relevant to the data; these are: **Home**, **Browse the Catalogue**, **Analyze the Data**, **Get the Data** and **Survey 2017**. Whereas the **Output**, **Documentation**, **FAQ**, **User Feedback**, **Imprint** and **Login** tabs on right-hand side provide contextual information about the project.

**Home.** This page provides an overview of the project, a collapsible news section used to publish significant updates and a Twitter feed of the Catalogue’s dissemination channel (#digEdsCat hashtag).
3.3. The Web Application

**Browse the Catalogue.** This page (Figure 3.11) provides the default table display of the data with the total count of digital editions present in the Catalogue at any one time.

*Figure 3.11:* The web application’s *Browse the Catalogue* page. Users can either serendipitously browse the digital editions or look for specific projects by using the faceted search provided (April 2017 snapshot).

The page can also be reached by clicking on the navy blue ‘explore’ button placed on the application’s home page. Beneath the page title, a single line of text is provided to build user trust as it publishes the time-stamp of the last data update. To the left side of the screen, a faceted search functionality allows users to locate specific digital editions; each filter or option in the faceted search corresponds to a feature in the Catalogue. Alternatively, the bottom pagination function can be used for a more serendipitous browsing experience (Figure 3.12).

*Figure 3.12:* Users who are not looking for a particular digital edition can browse the Catalogue with the help of a pagination function at the bottom of the screen, allowing them to jump between pages of data (April 2017 snapshot).

To view a digital edition from the *Browse the Catalogue* table list the user must click on the name of the digital edition in question. Figure 3.13 illustrates the layout of digital edition sheet in the Catalogue: the left portion
of the sheet lists a selection of project details from the digEds_cat.csv file, while the right side hosts a zoomable map displaying the location markers of the institutions involved. These markers are extracted from the institutions_places_enriched.csv file, which collects the geographical coordinates of each producing institution.\textsuperscript{23} In line with Linked Open Data (LOD) best practice recommendations, each marker is enriched with stable GeoNames and Deutsche Nationalbibliothek (DNB) URIs in order to link the Catalogue data to existing open data-sets on the web and thus foster knowledge discovery.\textsuperscript{24}

The Catalogue further integrates its data with the RIDE Review Journal for Scholarly Digital Editions and Resources to link projects to written reviews, if any,\textsuperscript{25} and with Patrick Sahle’s Catalog of Digital Scholarly Editions to indicate whether a project is also featured there.

In the sheet header, blue arrows enclose the project name. These allow users to move from project to project sequentially. Beneath the project name, a drop-down Download button reveals RDF and Notation 3 renditions of the data, and a black i information button optionally expands the written record to display brief descriptions of each feature in the list (Figure 3.14).

Below the map, a report button directs the user to the GitHub issue page to communicate a problem with the project, and a share button at the bottom of the sheet can be used to circulate the project via Twitter, Facebook and Google Plus.

The URL of the digital edition can be opened in a new browser tab so as to not lose sight of the project’s overview sheet. As discussed in section 3.2.2 of this chapter, some fields in the Catalogue contain decimal numbers as values to measure the degree of compliance of a project to a particular feature: 0 stands

\footnotesize
\textsuperscript{23}The geographical coordinates of funding bodies and agencies are not collected in the Catalogue yet.


\textsuperscript{25}The RIDE Journal is accessible at: http://ride.i-d-e.de/ (Accessed: 17 May 2017).
3.3. The Web Application

Figure 3.13: Example digital edition sheet in the Catalogue, with project details on the left and a map pointing to the involved institutions on the right. Clicking on the location marker brings up a pop-up window with more information about the institution (April 2017 snapshot).

<table>
<thead>
<tr>
<th>URL</th>
<th>The project URL</th>
<th><a href="http://claviusontheweb.it/">http://claviusontheweb.it/</a></th>
</tr>
</thead>
<tbody>
<tr>
<td>Historical Period</td>
<td>Historical period to which the source documents belong.</td>
<td>Early Modern</td>
</tr>
<tr>
<td>Scholarly</td>
<td>Scholarly - An edition must have critical components - a pure facsimile is not an edition, a digital library is not an edition.</td>
<td>yes</td>
</tr>
<tr>
<td>Digital</td>
<td>Scholarly - A digital edition cannot be converted to a printed edition without substantial loss of content or functionality - vice versa. A non-scholarly periodical edition is not a Scholarly Digital Edition (but it may evolve into a Scholarly Digital Edition through new content or functionalities).</td>
<td>yes</td>
</tr>
</tbody>
</table>

Figure 3.14: The i information button, when pressed, displays the descriptions of each feature in a collapsible column placed between the feature column and the value column (second column in this image) (April 2017 snapshot).
for ‘no compliance’, 0.5 for ‘somewhat/partially compliant’, 1 for ‘fully compliant’. The use of numbers makes it easier to run calculations and statistical analyses over the full data-set. However, for readability purposes the web application is set to transform these values into text, whereby 0 becomes ‘no’, 0.5 becomes ‘partial/somewhat’ and 1 becomes ‘yes’. Similarly, the web application translates the 0 value of the ‘Current availability’ feature into a red banner (Figure 3.16) to inform users of an ostensibly discontinued project. The Catalogue does not make a distinction between a temporary and permanent project shutdown.

Figure 3.16: A red banner above the sheet header of the Saint-Aubins Book of Arses digital edition informs users that the project is no longer available under the URL originally provided (April 2017 snapshot).

If the project resurfaces under a new URL, the ‘Current availability’ value is changed to 1 making the red banner disappear from its corresponding Catalogue sheet.
Analyze the Data. This menu item provides a drop-down list of two sub-pages: Select Charts and A Map. As of April 2017, the Catalogue contains 49 columns or features and an ever-increasing number of rows, showcasing a large data-set, which can only be fully and adequately studied with the help of visualisation. In the early days of the Catalogue, Google Fusion Tables was used to place all digital editions onto a map of the world. A number of pie and bar charts were also used but these had to be manually updated with every new data change (Franzini et al., 2016). Today, Catalogue visualisations are generated and updated automatically through the project’s integration with the open-source Highcharts JavaScript charting framework.\footnote{For more information about Highcharts, visit: http://www.highcharts.com/ (Accessed: 8 April 2017).} Available charts can be viewed by selecting the Select Charts entry in the Analyze the Data drop-down menu (Figure 3.17), and can be downloaded in different formats (PNG image, JPEG image, PDF document, SVG vector image) and/or printed. A menu icon in the top right corner of any selected chart reveals the download options (Figure 3.18).

<table>
<thead>
<tr>
<th>Property to visualize</th>
<th>What will it show</th>
<th>Type of visualisation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Period</td>
<td>Historical periods covered</td>
<td>bar chart</td>
</tr>
<tr>
<td>Scholarly</td>
<td>Critical vs. non-critical projects</td>
<td>pie chart</td>
</tr>
<tr>
<td>Digital</td>
<td>How digital projects are</td>
<td>line chart</td>
</tr>
<tr>
<td>Edition</td>
<td>Editions vs. Archives, Databases, Collections, etc.</td>
<td></td>
</tr>
<tr>
<td>Language(s) of the source text(s)</td>
<td>Language(s) in which the source text(s) are written</td>
<td></td>
</tr>
<tr>
<td>Writing support</td>
<td>Types of writing support</td>
<td></td>
</tr>
</tbody>
</table>

Figure 3.17: The user can choose one of three types of visualisation (bar, pie or line chart) for the list of features provided (April 2017 snapshot).

Figure 3.18: A menu icon in the top right corner of every Catalogue chart opens up download options in a pop-up window.
supported by the Catalogue. For instance the *Editions per Country* charts arrange producing institutions by country (Figure 3.19).

![Figure 3.19](image)

**Figure 3.19**: Two different views, a bar and a pie chart, of the number of digital editions produced per country. Hovering over the Italian slice of the pie brings up a pop-up window with the exact count of digital editions produced in Italy (April 2017 snapshot).

In the charts above, the reader will notice a shortage of Asian, African and Latin American editions. Reasons for this large gap in the data result from my own knowledge of languages. Browser plugins to automatically translate webpages exist but I hesitate to use them, as any machine-translation errors would go unnoticed and possibly lead to incorrect cataloguing (Figure 3.20). This is an example of where user contributions become necessary to help paint a global, rather than a Western-centric, picture of digital edition initiatives.

---

3.3. The Web Application

Figure 3.20: Google Chrome’s web translation plugin automatically detects the language of a particular page and offers the user the option to translate the text into another language. The plugin relies on the machine translation algorithms of Google Translate.

It should be noted, however, that Asian, African and Latin American colleagues do not appear to be using mainstream English Digital Humanities communication channels to advertise and promote these projects, thus hampering discoverability. Indeed, in the case of the Japanese Association for Digital Humanities, the official website does not exhibit the efforts conducted by Japanese institutions.28 We should not, however, exclude the possibility that digital editions are mostly a Western reality. A study by the Oxford Internet Institute (OII) reveals a correlation between wealth and data openness; data openness, OII explains, is partly dependent on Internet penetration.29 Countries affected by limited access to the Internet in Asia, Africa and Latin America appear at the bottom of the OII charts. While there might not be an obvious correlation between production of digital editions, openness and wealth, the data collected for the Catalogue thus far points in this direction. Dissemination is part of the European Digital Agenda, which provides a funding infrastructure to promote growth in the educational, cultural and commercial sectors, helping Europe build a competitive research and innovation profile.30 While North American investment in the information and communications technology (ICT) infrastructure is much higher than that of Europe, both continents are at the forefront of the educational sector.31 Finally, it is also important to consider that some of the major Digital Hu-

29 Emily Badger, Why the wealthiest countries are also the most open with their data (2014), at http://www.washingtonpost.com/blogs/wonkyblog/wp/2014/03/14/why-the-wealthiest-countries-are-also-the-most-open-with-their-data/ (Accessed: 18 March 2014).
84 Chapter 3: A Catalogue of Digital Editions

...manities associations and funding agencies are based in the US, UK and in Germany.32

Another numerical disparity one can observe from the Catalogue is in the representation of historical periods. The Catalogue broadly categorises source texts into *Antiquity [700 BC-500 AD], Middle Ages [500-1500], Early Modern [1500-1789], Long Nineteenth Century [1789-1914], Modern [1914-1965] and Contemporary [1965-today].* If a digital edition covers two categories, both periods are added and separated with a semicolon. Figure 3.21 shows that, as of April 2017, the Middle Ages are the most covered historical period in the Catalogue with some 70 digital editions.

![Fig 3.21](image)

**Figure 3.21:** Coverage of historical periods in the Catalogue (April 2017 snapshot).

The second most represented category is the Early Modern period and only one edition in the Catalogue publishes a contemporary text. The balance between digital editions of ancient (i.e. Antiquity and Middle Ages) and modern (i.e. Early Modern, Long Nineteenth Century, Modern and Contemporary) texts is more even now than in the first version of the Catalogue. In 2013–2014, digital editions of modern texts outnumbered those of ancient texts. The motivations given at the time to explain this imbalance included sampling error (e.g., data collection

---

32 Among many others, the Association for Computers and the Humanities (ACH); the Alliance of Digital Humanities Organizations (ADHO); the European Association for Digital Humanities (formerly ALLC); DHCommons; Digital Humanities Now; Humanities, Arts, Science, and Technology Advanced Collaboratory (HASTAC) and The Humanities and Technology Camp (THATCamp).
at that point just so happened to be more representative of modern texts) and the hypothetical existence of a higher number of modern texts (Franzini et al., 2016). The latter was partly informed by an article in which John Lavagnino wrote:

In the last few decades, many textual scholars have come to believe that classical texts and modern texts have very different kinds of textual problems and constitute different kinds of literary works. Texts from classical antiquity have great textual problems: any manuscript that has survived to our day of such texts is the product of a long sequence of copyings and recopyings, so that it’s likely to be full of errors in transmission that need to be corrected. These are errors on such a scale that the works are often simply unreadable without editorial correction. But for modern texts, the body of surviving evidence is very different. For texts circulated since the invention of movable type, and particularly for texts written since the seventeenth and eighteenth centuries, the problem of mistransmission is less and less imposing. The texts have been copied only a few times prior to the creation of our sources, rather than many times, and we often have many more of the sources, sometimes going back to the author’s own drafts. Error will always be present, and is still sometimes a great problem, but it ceases to be the central problem. What we have for many modern works is not a shortage of reliable information, but an excess: often there is far too much textual information to include in any printed edition. We find, for example, cases in which a writer made extensive revisions over a span of many years, so that there may be a number of versions that were all produced by the same person and that all have good claims to our attention; but which one should be the text that a scholarly edition prints? (Lavagnino, 1995, p. 111)

Of course, a low number of digital editions of contemporary texts is also be explained by the copyright restrictions imposed on manuscripts from the 20th and
21st centuries.

In 2012, Paolo Monella echoed Lavagnino’s observations with an article entitled ‘Why are there no comprehensively Digital Scholarly Editions of Classical Texts’, in which he argued that their absence could be explained by the fact that classicists do not need digital editions to conduct their research.\(^{33}\)

As of April 2017, that former imbalance in the Catalogue is leveled out with 105 digital editions jointly belonging to Antiquity and the Middle Ages against the remaining 127. This picture might still be symptomatic of sampling error and it attests to the need for more data in order to draw statistically significant inferences.

Most scholars working in the field of digital scholarly editing will be aware of Creative Commons Licenses and of their role in the fair dissemination and reuse of creative works.\(^{34}\) In 2014, Europeana for Research and Tourism published a booklet of Policy Recommendations in which it encouraged the adoption of Creative Commons Licenses to enable open access to digital cultural heritage (for Research and Tourism, 2014, p. 7). In the same year, the Times Higher Education reported that a Research Information Network study had found that Open Access articles published in the science journal Nature Communications in the first half of 2013 were cited twice as often as those accessible through subscription (Jump, 2014). Yet, despite this push, as the chart in Figure 3.22 shows us almost 150 digital editions in the Catalogue do not employ any form of Creative Commons License. The licensing of the remaining projects is either proprietary, only partially Creative Commons or unclear.

Again, one could question the representativeness of the data sample, but even if that number were double it would still raise some concern.

Advocates of the Text Encoding Initiative (TEI) would be glad to learn that 118 digital editions in the Catalogue encode their texts in TEI-compliant XML. How-

\(^{33}\)The full article can be read here: http://www1.unipa.it/paolo.monella/lincei/why.html (Accessed: 13 April 2017).

\(^{34}\)See http://creativecommons.org/ (Accessed: 24 April 2013).
3.3. The Web Application

Figure 3.22: Almost 150 digital editions in the Catalogue do not make their contents available in open access (April 2017 snapshot).

However, there are also many projects that do not make use of any XML, be that TEI or custom-made, preferring instead to publish plain text or HTML code. Furthermore, a small number of digital editions are unclear about their treatment of the texts and one forthcoming edition has not made its intentions known yet (hence, ‘N/A’).

Figure 3.23: The largest group of digital editions in the Catalogue encode their electronic texts in TEI-compliant XML (April 2017 snapshot).

In the questionnaire described at the outset, some creators expressed a preference for custom XML and/or alternative technologies over XML-TEI because TEI has to be learnt and does not always meet the encoding needs of the editor. Truthfully, however, concerns regarding the suitability of TEI for the encoding
of all forms of manuscript culture are already being addressed by the TEI development community through the establishment of Special Interest Groups (SIGs), whose objective it is to develop encoding schemata tailored to specific primary sources. Yet, as Chapter 4 of this thesis reports, many users of digital editions also expect Plain Text, PDF, Microsoft Word and HTML formats.

On final example visualisation of Catalogue data concerns digital imaging. The digitisation of GLAM (Gallery, Library, Archive and Museum) holdings makes available high-quality images, often for free use. Imaging technology is being developed to help manipulate, annotate and integrate images into various systems, including digital editions. Notable examples are the previously mentioned International Image Interoperability Framework (IIIF), currently adopted by large memory institutions such as the Library of Congress and the Biblioteca Vaticana, and Mirador, deployed by the Swiss e-codices Virtual Manuscript Library. As Figure 3.24 shows, as of April 2017, 128 digital editions in the Catalogue provide access to digital images of the primary source(s). The Catalogue does not yet distinguish between accessible and downloadable and/or reusable images but does indicate the presence/absence of zooming and editing functionality, as well as text-image linking.

Figure 3.24: The majority of digital editions currently in the Catalogue provide digital images of the primary source(s) (April 2017 snapshot).

---

The second entry in the Analyze the Data dropdown menu is a map view of all digital editions in the Catalogue. The searchable Map locates institutions contained in the institutions_places_enriched.csv file with coloured markers; each marker, if clicked, lists all digital editions relating to the corresponding institution; to view the digital editions, one need only click on the chosen project to be directed to its Catalogue sheet (Figure 3.25). The map comes with a search functionality to filter the markers on the map.

Figure 3.25: Map visualisation of the institutions recorded in the Catalogue. The only marker details a single digital edition from the Indian Ministry of Culture; when clicking on the digital edition, the user is directed to the project’s sheet in the Catalogue (April 2017 snapshot).

Get the Data. This menu item provides a drop-down list of five sub-pages, JSON, Download institution and places (.csv), BibTeX, N3 and RDF/XML. JSON is the Application Programming Interface (API) of the Catalogue, a set of rules that describes the interactions between the different technical components of
the web application. Download institution and places (.csv) provides a copy of the institutions_places_enriched.csv file. The BibTeX data format allows users to import the digital editions as bibliographic entries into a bibliographic manager, such as Zotero.\footnote{For more information about Zotero, see \url{https://www.zotero.org} (Accessed: 20 September 2017).} Finally, the Notation 3 (N3) and Resource Description Framework (RDF/XML) weave the Catalogue data into Linked Open Data vocabularies, such as Friend of a Friend (FOAF) and the Data Catalog Vocabulary (DCAT).\footnote{For more information about FOAF, see \url{http://xmlns.com/foaf/spec/} (Accessed: 15 September 2017). For more information about DCAT, see \url{https://www.w3.org/TR/vocab-dcat/} (Accessed: 15 September 2017).}

**Survey 2017.** This page will provide the results of the survey described in Chapter 4 after the publication of a related forthcoming article.\footnote{The expected publication time is between late 2018 and early 2019.}

**Output.** This page gathers all project presentations, posters and publications, as well as a Refbacks section to keep a record of the Catalogue’s presence in other resources.

**Documentation.** This page documents the project as it progresses. It aims at providing sufficiently detailed information to guide the user through the various components of the project.

**FAQ.** This page collects frequently asked questions.

**User Feedback.** This page collects testimonials and comments from users as a means of sharing experiences and attracting new users. User feedback received since the beginning of the collaboration in July 2016 has been encouraging. Some users have requested new features, such as the provision of permalinks (Figure 3.26), information about the size of the text(s) in tokens and information concerning data pre-processing steps taken to prepare the text for editing.

Comments on the website or the user interface of the web application led to the improvement of the pagination system and to the introduction of a boxed help-
3.3. The Web Application

Figure 3.26: A user request made via the web application’s GitHub repository asked for the provision of permalinks for all projects in the Catalogue. While the application currently provides URIs, these are not stable identifiers yet.

text to clarify the difference between page navigation with the browser’s and the Catalogue’s back/next arrows (Figure 3.27).

Figure 3.27: A removable blue information box explains how to use previous/next arrows to navigate the Catalogue (April 2017 snapshot).

Imprint. This page is a legally mandated statement of project ownership typical of German-speaking countries.

Login. This page gives the project team access to administrative functionality, including news management and data synchronisation.

3.3.2 Dissemination

Updates are published under the News section of the project’s home page on a monthly basis. The social dissemination channels used to spread these updates are Twitter (#digEdsCat hashtag) and ResearchGate. Furthermore, the Catalogue is being added to any relevant resources that might help it to gain wider exposure.42

42For example, the European Association for Digital Humanities (EADH) project list, which is available at: https://eadh.org/projects (Accessed: 19 April 2017).
3.3.3 User analytics

The ACDH makes use of the open analytics platform Piwik to monitor visits to the web application.\textsuperscript{43} For every visit to the Catalogue, Piwik records the provenance (country, operating system and IP address), the time of access, the duration of the visit, the specific pages visited and the referral type (i.e. whether the user accessed the Catalogue directly or reached it via another website). Figure 3.28 gives an overview of the Catalogue’s Piwik landing page. The screen is divided into two main vertical areas: a left sidebar contains the primary menu, and a main area accommodates multiple windows or widgets with data summaries.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figures/piwik_landing.png}
\caption{The Catalogue’s Piwik analytics landing page. The primary menu is located in the leftmost side-bar and the majority of screen space is given to the data widgets (September 2017 snapshot).}
\end{figure}

The page displayed in Figure 3.28 can be customised: administrators can add more widgets for a more detailed overview of data facets, and can enlarge or rearrange widgets. Within the widgets, charts can be changed to display a particular metric. As Figure 3.29 shows, the \textit{Visits Over Time} widget offers numerous metric possibilities, while the \textit{Visitor Map} provides views for each continent.

\textsuperscript{43}In 2018 Piwik was rebranded to Matomo. For more information about Piwik/Matomo, visit: https://matomo.org/ (Accessed: 10 April 2018).
Figure 3.29: Data widgets can be customised to view specific layers of information (September 2017 snapshot).

Piwik also gives a full visitor profile, including time of access, IP address, country of access, details regarding the browser, operating system and device used to access the resource, and the specific page(s) visited (Figure 3.30). Piwik also assigns an ID to visitors in order to distinguish returning from unique visitors.

Piwik data can be exported to multiple formats, including CSV, TSV, XML, JSON, PHP, RSS and as an image, giving users the possibility to reuse and incorporate this information in other resources.
Figure 3.30: The Piwik Visitor Log lists visitors of the Catalogue. Moving the cursor over the three middle icons will bring up more information in a black pop-up window. Alternatively, the user can click on View visitor profile in the top right corner to enlarge the visitor card (April 2017 snapshot).

The amount of information Piwik collects helps measure the impact of the monitored resource. For example, on 18 April 2017 the Catalogue was accessed from a service at the University of Regensburg in Germany (Figure 3.31). Upon closer inspection, I learnt that the Catalogue had been syndicated by the Datenbank-Infosystem (DBIS), a German information service established at the University of Regensburg to collect scientific databases for use in hundreds of libraries across the nation. As of April 2017, DBIS makes the Catalogue available to over 300 German libraries.

Figure 3.31: On 18 April 2017 Piwik revealed that the Catalogue had been included in the German Datenbank-Infosystem (DBIS) (April 2017 snapshot).

Piwik proves to be effective in providing statistics about how users access the Catalogue and which pages they spend more time on. Such information can be used to improve the dissemination strategy of the Catalogue by studying and targeting specific communication channels. Additionally, it can be used to revise the content of the Catalogue website to retain user interest and engagement across the entire resource. However, as Piwik does not measure usefulness and specific use, a dedicated study is needed in order to better understand how the

---

3.4. Sustainability

The continued maintenance of the Catalogue of Digital Editions currently relies entirely on the commitment and employment status of the team (Peter Andorfer and Ksenia Zaytseva of the ACDH, and myself). I am personally committed to maintaining the data component of the project, so should the web application be discontinued, the data itself would not be affected. The project is regularly backed-up on Austrian servers but there is no indication at this time of any hosting restrictions to the Catalogue. Nevertheless, to maintain the project in the long term, continuous efforts must be injected into data addition and curation, into the development of the application, into dissemination and thorough documentation. The more useful and used this resource becomes, the stronger the argument becomes for long-term sustainability.
3.5 Future work

Despite the success, more work remains to be done to turn the Catalogue of Digital Editions into an ever-useful resource.\textsuperscript{45}

In the short term, upcoming tasks include the stable release of the web application, its adaptation to the Accessibility Review WCAG 2.0 (Level AA) Guidelines and the publication of technical documentation.\textsuperscript{46} Minor tasks and features are listed in the Issues lists of the project’s GitHub repositories. In the longer term, the team will seek further collaborations and partnerships to, for instance, better adhere to library cataloguing standards and broaden the Catalogue’s Linked Open Data radius. Once the manuscript transcription is complete, a survey will be prepared to gather feedback on the usefulness and use of this project.

3.6 Conclusion

This chapter describes the creation and ongoing development of the Catalogue of Digital Editions. The aim of the project is to catalogue extant digital textual editions in an effort to survey past and current diversity of practice in the creation of digital editions.

In 2016, the development of a web application at the Austrian Centre for Digital Humanities increased the project’s visibility and usability. In early 2017, the Catalogue was included in the German national Datenbank-Infosystem (DBIS) for use in over 300 libraries, inherently becoming a trusted scientific database. Since 2016, the number of user contributions has been proportionally higher than the 2012-2015 period, suggesting that the visibility awarded by the web application increases user engagement. Starting off as a personal tool of research,

\textsuperscript{45}Updates to the project since this chapter was written include the addition of network visualisations and of a SPARQL interface to help users query the Linked Open Data network of the Catalogue of Digital Editions. The Catalogue was also indexed by OpenAIRE at: https://www.openaire.eu/search/dataset?datasetId=r37b0ad88687:c2ad7b522d5abc748a1c6bb807aeb (Accessed: 25 April 2018).

the number of entries submitted by users since its publication (on an almost monthly basis) suggest that the Catalogue is seen by the community as a reference portal for the dissemination of digital editions. This interest has motivated the team behind the Catalogue to look for external help to help maintain the project.\footnote{In 2018, the team was joined by an enthusiastic volunteer from Hungary, who is responsible for many of the Hungarian projects listed in the Catalogue. I am looking for Bachelor’s and Master’s students to help with data entry and curation.}

The evidence gathered and centralised in this unique searchable registry facilitates the study of the evolution of digital editing. It is only by thoroughly cataloguing and analysing the hundreds of digital editions that now exist that we can understand and question the scope of the field, spot technical and procedural trends, and make recommendations as to how best to build digital editions that provide the information required by users. 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)
To the best of my knowledge, no definitive estimation of the total number of digital editions produced to this day exists. Thanks to the Catalogue of Digital Editions, that question may soon find an answer.
Chapter 4

User expectations of digital editions

Despite the large number of digital editions being created, little is known about the 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. The study described in this chapter follows on from the few existing efforts in this relatively unexplored area and seeks to answer the following research question: what are the user expectations of users of digital editions in the Arts and Humanities and how do these correlate with existing digital editions?

This chapter compares the findings of Chapter 3 with the results of a user survey that identified the important and salient features that the Digital Humanities community most desires from digital editions. It is the first comparative analysis of the differences between what the 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.
Chapter 4: User expectations of digital editions

4.1 Users, use, usage, usefulness and usability

Before looking at users and use of digital editions, it is first necessary to understand and distinguish between the terms users, use, usage, usefulness and usability. The Oxford English Dictionary University (2011) defines a user as:

A person who has or makes use of a thing, esp. regularly; a person who employs or practices something. (User, n1.a)

And defines the noun use as:

The action of consuming something […] esp. on a regular or habitual basis. (Use, n1.1.b)

This meaning of use is only subtly distinct from usage, which the OED defines as:

An established or recognized mode of procedure, action, or conduct; a custom or practice; spec. one which has force of law. (Usage, n1.1.a)

It defines usefulness as:

The state or condition of being useful or serviceable; utility, serviceableness. (Usefulness, n2)

And usability as:

The fact or quality of being usable. (Usability, n1)

Whereby the adjective usable is employed to denote something:

That can be used; that can be readily put to practical use. (Usable, n1)

Based on these definitions, use and usage are the dimensions that more fittingly describe the user-side of digital editions, whereas usefulness and usability are more pertinent to the creators. To lay it out as a paradigm:

makers put users in a position to consume (=use) the edition, ideally
more than once (usage), by offering practical (usable) means of enjoying its convenience (usefulness).

4.1.1 Users

In the digital realm, users are traditionally categorised into active users and passive users. Many definitions of active and passive users exist and often depend on the domain within which they are studied. For instance, human-computer interaction speaks of ‘active process operators’ and ‘passive process operators’, the difference lying within the level of predominance of monitoring tasks (Persson et al., 2001). Another example can be found in Internet Culture, for instance, where the 1% rule discriminates between creators, contributors and lurkers (Nielsen, 2006).1 These and other definitions of user types can all be reduced to the fundamental distinction between those who supply, the active, and those who consume, the passive.

Passive users are the prevalent type of users in the real world (Montague and Xu, 2012, p. 703). In the field of digital editions, we might understand passive users as those who interact with a digital edition for browsing purposes. This is not to say that passive users do not play a role in the creation of digital editions; on the contrary, passive users can help shape a digital edition by communicating problems and requirements to its creators.

Active users, on the other hand, can be understood as those users who contribute to, and more visibly engage with, the production of a digital edition by (where possible) reusing its content, developing the underlying technology or by supplying data (these users correspond to the advanced users described by (Dillen and Neyt, 2016)). Examples are social or crowdsourced digital editions, such as the Devonshire MS (BL Add 17,492) project or the Bentham Project.2

Whether active or passive, users are driven to use a resource based on their trust

---

1 Also known as the 1-9-90 rule or the 90-9-1 principle or the 89:10:1 ratio.
in its content, which can be secured, for instance, by maintaining high academic
standards and by being clear as to the motivation, the intention and methodo-
logy of the digital edition. Trust gained translates into higher user satisfaction
and engagement (Inbar and Tractinsky, 2011, pp. 249-250). A trustworthy digital
edition is more likely, for example, to secure follow-up funding, to be employed
as a teaching aid or to be reused in other initiatives.3

4.1.2 Use

Dillen and Neyt (2016) identify ‘simple browsing functionalities’, ‘research the
materials the Digital Scholarly Edition has to offer’ and the ‘re-use of the Digital
Scholarly Edition’s data’ as the three types of uses creators of digital editions can
expect (Dillen and Neyt, 2016, p. 3). Further, Kelly reports that users typically in-
teract with a digital edition for research purposes, personal curiosity or to com-
plete assignments, and that the two most popular uses of digital editions are the
study of digital images and the extraction of text for computational analysis. Ac-
cording to Kelly’s study, users consider the text to be the most important feature
in a digital edition and the tablet functionality most sought after is text annota-
tion (Kelly, 2015, pp. 132, 133, 136). These findings resonate with the results of
the survey described here (see section 4.3.1).

4.1.3 Usefulness

Little evidence attests to the usefulness of digital editions. In project proposals,
creators typically attribute their decision to undertake digital edition projects
to the affordances and advantages of the digital medium over print, in that web
presence facilitates the access, study and dissemination of source materials in
ways never before possible (Dillen and Neyt, 2016, p. 1) (Terras, 2012a). Yet,
while acknowledging the usefulness and envisaging ways in which digital edi-
tions enhance print counterparts, creators of digital editions rarely take the time
to study how their predictions actually translate into practice: does the user ex-

3This research does not take into account users with computer anxiety. For more informa-
tion about the topic, see (Jones, 2010).
Users, use, usage, usefulness and usability

Experience live up to expectation? Has the project been successful in providing innovative and beneficial means of approaching existing or previously inaccessible materials? These questions cannot be answered through webometric techniques or log-files but require direct communication with users in the form of feedback forms, surveys or working groups.4

4.1.4 Usability

The discussion in Chapter 2 about the usability of digital editions focussed on the differences between the print and digital medium. There are advantages and disadvantages to each but thanks to web technologies digital editions lend themselves better to targeting content at specific audiences or to presenting different views of the same text.5 Moreover, digital editions can enrich users’ experience with additional materials drawn from other digital resources or with visualisations. For example, the translations of Latin terminology in the German transcriptions of the Diary of Calvinist Prince Christian II of Anhalt-Bernburg can be displayed on demand.6 It is easy to see how such possibilities would support a wider or non-academic audience to peruse the contents of the digital edition.

Another key aspect of the usability of a digital resource is responsiveness, that is, the ability of a system to complete a task within a given time-frame. In the case of digital editions, responsiveness includes the ability to use the resource across various devices, namely desktop computers, smart phones, connected TVs, tablets and retinal displays (Gardner, 2011). Indeed, while the uptake of hand-held devices is driving the development of flexible technologies, digital editions of texts do not appear to be adapting to mobile access (see section 4.3.1). Kelly also

---

4 A description of the peculiarities of webometric analysis is beyond the scope of this discussion. However, for an example study see Gooding (2016).
5 For example, the Ancient Inscriptions of the Northern Black Sea project (IOSPE) provides edited, diplomatic and XML views of each inscription. See for example: http://iospe.kcl.ac.uk/5.7.html (Accessed: 27 August 2016).
reveals that access to digital editions from tablets is not common:

Only 30% [of interviewees] had previously attempted to use a digital edition on a tablet. This low attempt rate may be due to a combination of their expectations that the user interface wouldn’t work well and perhaps also because tablet usage currently fits better in the respondents’ leisure/reading patterns than with their working/research patterns. Our interview discussions indicated that this is the case, and that users sometimes have other software that they need to access while using the editions, which is only available to them on their desktop machines. (Kelly, 2015, p. 135)

This point brings us to one final consideration regarding usability. Usability concerns consumers as much as it does contributors. The usability of digital editions for consumers can be understood as the trade-off between using the highly and accurately annotated data provided while still being able to read the texts. Usability for the contributors is the trade-off between the quantity of annotated data and the quality of the annotations. So, the question designers of digital editions might ask themselves is: what is the balance we need to strike in order to maximise effort and usability?

### 4.2 Related work: Users studies of digital editions

In 2008, the LAIRAH (Log analysis of Internet Resources in the Arts and the Humanities) initiative analysed twenty-one digital resources and found that:

P[principal]I[nvestigator]s […] infer[red] user requirements from their own behaviour. (Warwick et al., 2008, p. 390)

Indeed, that same year, an article entitled The Inscriptions of Aphrodisias as electronic publication: A user’s perspective and a proposed paradigm inferred the use

---

of an electronic resource, the Inscriptions of Aphrodisias project, based on its author’s own experience (Bodard, 2008).

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, Porter published an article entitled Medievalists and the Scholarly Digital Edition, which summarised the results of two surveys 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). Porter’s 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 (Porter, 2013, 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, 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 (Porter, 2013, 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, Porter 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” (Porter, 2013, p. 14).

More recent monographs on digital scholarly editing, such as Digital Critical Editions by Apollon et al. 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 (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. Interfaces of digital editions were also the central theme of the 2016 conference *Digital Scholarly Editions as Interfaces* organised by the *Digital Scholarly Editions Initial Training Network* (DiXiT). In the same year, and as previously mentioned, (Dillen and Neyt, 2016, p. 3) 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 the DSE provides, or to perform functionalities the DSE does not (yet) offer (and publish the results).

---


10 For more information about DiXiT, see: http://dixit.uni-koeln.de/ (Accessed: 3 July 2017).
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. However, no overarching analysis has been done on community investment in digital resources, as compared to user desires. 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 a 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 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). In *Studying Users in Digital Humanities*, Warwick stresses the importance of examining user needs:

---

11Available at: [http://ride.i-d-e.de/](http://ride.i-d-e.de/) (Accessed: 3 July 2017).
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 chapter reports on the first initiative to bring together a varied group of users of digital editions with a view to identifying needs and expectations.

4.3 Method

“good idea this survey!”

To answer this study’s research question (what are the expectations of digital editions in the Arts and Humanities and how do these correlate with existing digital editions?), on 30th March 2017 I 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 primarily targeted at the Digital Humanities community, the field I identify with and within which many digital...

---

12Comment left by a participant of this survey (in reply to Question 20 - see section 4.3.1).
13The survey was distributed via my Twitter account, and to the following mailing lists and Facebook groups:
   Associazione per l’Informatica Umanistica e la Cultura Digitale (AIUCD) at http://lists.lists.digitalhumanities.org/mailman/listinfo/aiucd-l;
   Corpora at http://clu.uni.no/icame/corpora/sub.html;
   Digital Humanities im deutschsprachigen raum (DHd) at http://dig-hum.de/dhd-mailingliste;
   Digital Classicist at https://www.jiscmail.ac.uk/cgi-bin/webadmin?A0=DIGITALCLASSICIST;
   Digital Medievalist at http://listserv.uleth.ca/mailman/listinfo/dm-l;
   Humanist at 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/;
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 (Horner, 2011, p. 956). 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,\(^\text{14}\) 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).

### 4.3.1 Survey design

Surveys are often imperfect and introduce a number of issues, including sampling bias, coverage or sampling error, nonresponse bias, self-selection bias and representativeness.

**Sampling bias.** Sampling bias occurs when a sample is collected in such a way that intended or relevant respondents might be left out (McCutcheon, 2011, p. 785). In the case of this web survey, the bias occurred by sending out the questionnaire to a selection of mailing lists and other communication channels, potentially excluding relevant respondents. By targeting only Digital Humanities Facebook group at [https://www.facebook.com/groups/49320313768/](https://www.facebook.com/groups/49320313768/); European Association for Digital Humanities (EADH) Facebook group at [https://www.facebook.com/groups/109971049335068/](https://www.facebook.com/groups/109971049335068/); UCL Centre for Digital Humanities Facebook group at [https://www.facebook.com/UCLDH/](https://www.facebook.com/UCLDH/) (All accessed: 3 July 2017). Facebook and Twitter posts were re-advertised by individuals in numerous other social spaces.

\(^{14}\)Early “offline” examples of digital editions such as the *Thesaurus Linguae Graecae* (TLG) were published on CDROM, but these are now rare (and often replaced, like the TLG, by an online version), difficult to access or expensive to purchase.
communication channels, the present survey introduced *sampling errors* (see below) as it likely excluded potential candidates from other fields of study. That said, exposing the survey to a wider and undefined frame population (for example, language and literary studies, social sciences, legal history and many others) would have led to scoping issues: which communities to include and which to exclude? As the first survey of its kind, it was deemed sufficient to begin by focussing on a community known to be active in digital scholarly editing. In future, however, the search should be extended to a wider audience.

**Sampling or coverage error.** As Couper notes, one of the major sources of error in web surveys is *sample or coverage error* or, as he puts it, “[…] the function of the mismatch between the frame population and the target population.” (Couper, 2000, p. 467). What this means is that the frame and target populations might be missing relevant survey candidates. In fact, users of digital textual editions may not belong to the Digital Humanities community or be familiar with Digital Humanities at all. Nevertheless, coverage error in this survey was to some extent decreased through the Twitter campaign of Professor Melissa Terras, whose large following covers a wide range of fields and institutions.15

> […] there is a misguided assumption behind many Web surveys that large samples necessarily mean more valid responses, or that sample size (or, more correctly, number of respondents) is the only element in sampling error. (Couper, 2000, pp. 472-473)

**Nonresponse.** This term is used to describe responses or questionnaires not returned. Merkle summarises nonresponse as follows:

> There are three basic types of survey nonresponse. The first is refusals, which occur when sampled individuals or households decline to participate. The second is noncontacts, when sampled individu-

---

15Professor Melissa Terras’s Twitter account is followed by some 13,000 users, covering a wide range of fields and institutions, and many of her followers multiplied the visibility of the survey by retweeting the announcement. See: https://twitter.com/melissaterras (Accessed: 23 April 2017).
als are never reached. The third type of nonresponse consists of situations in which the interviewer cannot communicate with the sampled person because of a language barrier or some mental or physical disability. Most nonresponse is the result of refusals and noncontacts. (Merkle, 2011, p. 532)

It is very likely, but almost impossible to verify, that this survey was affected by nonresponse either because the questionnaire did not reach compatible participants or because some of the participants who did receive it chose not to complete it.

Of the 348 responses stored by this survey, 130 were incomplete. One factor that may have contributed to incomplete responses is the “save and complete later” setting. Although the original intent of this setting was to give respondents the option to complete the survey in their own time, splitting completion between multiple sessions may have increased users’ likelihood of forgetting to submit a completed response. However, it is also possible that participants merely lost interest midway or were unable to participate due to computer hardware or software issues (Horner, 2011, p. 957).

Reminders to complete the survey can help reduce the nonresponse rate (Merkle, 2011, p. 532). In the case of the present survey, a single reminder was sent out three weeks into the four-week active period to encourage more responses. No additional reminders were sent in order to avoid negative reception:

[…] response rates may not be appreciably affected by larger numbers of reminder notices, and in fact a slight decrease among those receiving the largest number of reminders has been observed. (Cook et al., 2000, p. 831)

It is impossible to determine how many of the responses in the last week of the survey period were prompted by the reminder.
4.3. Method

Representativeness. 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 (Shoute 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, I will refrain from claiming representativeness, and acknowledge the qualitative nature of the results. While preliminary, the results of this survey are still valuable for those who are planning to build digital editions.

Confidentiality. Finally, as Horner reports, “[...] Web surveys may pose additional participant confidentiality issues”, which might deter participants from contributing their answers (Horner, 2011, p. 957). In this survey, demographic information about the participants, such as gender, age, ethnicity and religion, was not collected as it was not considered relevant for the study. The complete anonymity was announced in the survey’s advertising message and the software was set to not reveal any information about participants (see Appendix A.1). Instead, it assigned an ID to each participant, as shown in Figure 4.1.

In order to compare the results of the survey against the digital editions in the Catalogue of Digital Editions, the questions were modelled upon the 49 cataloguing features used in the Catalogue. This study excludes digital editions in the Catalogue whose ‘Digital’ value is 0 (0 = digitised; 1 = digital), which leaves...

---

Footnote: 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...
us with 242 projects eligible for the present analysis.\textsuperscript{17}

To encourage as many responses as possible, I 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 contained a mix of multiple choice and Likert scale questions\textsuperscript{18} 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 \textit{Opinio}, which is provided by my home institution for free.\textsuperscript{19} The survey broadly defined a digital

\begin{footnotesize}
\begin{itemize}
\item \textsuperscript{17}This analysis was done in May 2017 and thus ignores digital editions added to the Catalogue from June 2017 onward.
\item \textsuperscript{18}For more information about the Likert scale, see Allen and Seaman (2007).
\item \textsuperscript{19}Available at \url{https://www.ucl.ac.uk/isd/services/learning-teaching/}
\end{itemize}
\end{footnotesize}
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.

For every question in the survey, Opinio generated a chart containing the following statistical metrics:

- **Average**: the average score or arithmetic mean calculated from the responses given.

- **Median**: middle value of a number set. Given the survey’s values 1-5, the higher the median the more consensus among participants. For instance, given the values 10, 13, 17, 20 and 45 the median is 17, whereas the average is 21.

- **Minimum**: the lowest score given.

- **Maximum**: the highest score given.

- **Variance**: the average of the squared differences from the mean.

- **Standard deviation**: the amount of uncertainty around the mean value; in other words, standard deviation measures how spread out the numbers are. It is the square root of the variance. The closer the value of the standard deviation to the mean (small standard deviation), the less uncertainty and the more reliable the data/results; the more distant the value from the mean (large standard deviation), the more uncertainty and the lower the reliability of the data/results.

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, the closest initiative to this survey, recorded 149 completed responses over a period of four months and Porter’s 2002 and 2011 surveys recorded 43 and 169 completed responses respectively (Warwick et al., 2008, p. 385) (Porter, 2013).

In the next section, I 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 degrees of completeness do not make for a uniform analysis (however, these can be read in A.5). Median (of the five-point Likert Scale) and Standard Deviation (from the median) values will be hereafter referred to as $M$ and $SD$ respectively.

**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%); 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 manuscript culture 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 (see Appendix A.2 for the full list, which is too long for inclusion here), 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% on the Middle Ages. Other disciplines are as diverse as Agriculture, Architectural History and Anthropology, Bureaucracy, Caribbean Literature,
Celtic Studies, Children’s Literature, Conceptual History, Digital Archaeology, Engineering, Folklore, Food History, Historical Metrology, History of Daily Life, History of Exotic Animals in the Middle Ages, History of Ideas, History of Photography, Hittitology, Information Science, Journalism, Language Pedagogy, Law, Lexicology, Liturgical Medieval Chant, Medieval Legal History, Medieval Medicine, Military History, Musicology, Shakespeare Studies, Sinology, Social History of the Ottoman Middle East, Sociology of Culture and Science, Surgery and the History of Medicine, Television, Theatre and Visual Rhetoric. Those who identified as scientists include, among others, a computer scientist focussed on interaction design, a cybersecurity 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 they seek 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; and 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 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 (Sahle, 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 ex-
4.3. Method

Figure 4.5: Answers recorded for survey question 5.

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 used 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, while a completed project ran for sixteen years. 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 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.

---

20 The Edition Humboldt Digital project, for example, expects to reach completion in 2032. See: https://dig-ed-cat.acdh.oeaw.ac.at/editions/detail/251 (Accessed: 3 July 2017).

4.3. Method

Figure 4.6: Answers recorded for survey question 6.

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 their intended audience known helps users contextualise the data and better 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
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.

![Figure 4.7: Answers recorded for survey question 7.](image-url)
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.\textsuperscript{22} While reproducibility might not be essential to all disciplines, scholars argue that research should be sufficiently documented in order for it to be accepted by the community (Casadevall and Fang, 2010).\textsuperscript{23} 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 (Warwick et al., 2008, p. 391).

**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

\textsuperscript{22}Such as the Coursera Reproducible Research online course by Johns Hopkins University, available at https://www.coursera.org/learn/reproducible-research (Accessed: 3 July 2017).

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 a 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 technolo-
gies 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 pub-
ished 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. Disappoint-
ingly, these results corroborate the LAIRAH findings published over ten years
ago.

**Recommendation.** Given the low statistics currently discernible in the Cata-
logue and, conversely, the high rating of this question, creators must take more
care to incorporate comprehensive documentation. This should provide descript-
ive information about the project itself, including purpose, motivation, duration,
limitations, human and financial resources invested, and target audience. Ad-
ditionally, 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 pro-
ject back-end.
4.3. Method

Figure 4.8: Answers recorded for survey question 8.

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 in-
formation, 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 (which I was unable to access) 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 functionality 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 functionality 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
4.3. Method

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 (Fischer, 2017, pp. S284-S287).

**Figure 4.11:** Answers recorded for survey question 11.

**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, 2013), 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); 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.

24 Available at: https://chrome.google.com/webstore/detail/google-translate/aapbdbdomjkkjaonfhhkikfgjllcleb?hl=en-GB (Accessed: 3 July 2017).
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, 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.

---

25See, for example, Cohen (2010), Ramsay (2010) and Fitzpatrick (2010).
• 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.

• 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)*26*

**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

---

*26Excerpt 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 (Accessed: 3 July 2017).*
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.27

Figure 4.13: Answers recorded for survey question 13.

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).

27A comprehensive list of open source licences can be accessed at https://opensource.org/licenses (Accessed: 3 July 2017).
**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.

---

**Figure 4.14:** Answers recorded for survey question 14.
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.\(^{28}\) This survey question sought to understand users’ views on accessing digital editions from handheld devices.\(^{29}\) 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 casual 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*.\(^{30}\) 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,\(^{31}\) to champion inclusivity should become an integral part of the development process of a digital edition. The findings of a survey entitled “Inclusive


\(^{29}\) For a study on the dissemination of digital scholarly editions via mobile devices, see Kelly (2015).


\(^{31}\) Available at: [https://www.w3.org/](https://www.w3.org/) (Accessed: 3 July 2017).
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.\footnote{Available at: https://www.surveymonkey.com/r/MCDRMYY (Accessed: 18 July 2017).}

**Figure 4.15:** Answers recorded for survey question 15.

**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). The majority of participants (57%) rated the possibility of downloading and reusing the content of a digital edition as *very important*. One participant expressed no 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 here 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.

![Figure 4.16: Answers recorded for survey question 16.](image)
**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 often 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.\(^{33}\) 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 expectation 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 of the edited texts. This information, planned as a future addition to the Catalogue, would prove particularly useful for 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.

(White, 2017)

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)

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.3). 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

---

4.3. Method

statistical analysis. Download and reuse options could be better.

Digital editions mentioned by multiple respondents are the Perseus Digital Library (seven respondents),\textsuperscript{35} Electronic Beowulf (five respondents),\textsuperscript{36} Folger Digital Texts (four respondents),\textsuperscript{37} Online Froissart (three respondents),\textsuperscript{38} Loeb Classical Library (two respondents),\textsuperscript{39} Corpus Corporum (two respondents),\textsuperscript{40} e-codices (two respondents) and the Bayeux Tapestry (two respondents).\textsuperscript{41}

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 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.

\textsuperscript{35} Available at: http://www.perseus.tufts.edu/hopper/ (Accessed: 3 July 2017).
\textsuperscript{36} Available at: http://ebeowulf.uky.edu (Accessed: 3 July 2017); the project’s record in the Catalogue is available at https://dig-ed-cat.acdh.oeaw.ac.at/editions/detail/145 (Accessed: 3 July 2017).
\textsuperscript{37} Available at: http://www.folgerdigitaltexts.org (Accessed: 3 July 2017).
\textsuperscript{38} Available at: https://www.hrionline.ac.uk/onlinefroissart/ (Accessed: 3 July 2017); the project’s record in the Catalogue is available at https://dig-ed-cat.acdh.oeaw.ac.at/editions/detail/124 (Accessed: 3 July 2017).
\textsuperscript{39} Available at: https://www.loebclassics.com (Accessed: 3 July 2017).
\textsuperscript{40} Available at: http://mlat.uzh.ch/MLS/index.php?lang=0 (Accessed: 3 July 2017).
\textsuperscript{41} e-codices is available at: http://www.e-codices.unifr.ch/en (Accessed: 3 July 2017); the Bayeux Tapestry is available at: http://www.sd-editions.com/bayeux/zoom/#/facsimile%3DBayeux%26pane1%3D1 (Accessed: 3 July 2017).
The general impression one can glean from reading all of the answers given to this question is that users understand 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, I 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 A.4):

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.*  

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 that [sic] users can actually RTFM.42

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.

4.4 Discussion and recommendations

How do the digital editions in the Catalogue compare to the responses recorded by this survey? The overall correlation is best captured as a chart, as shown in Figure 4.19: 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 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 greater the distance between the red points and the tip of their respective 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.

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, I 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 from the start of the project. These studies should iterat-
4.5 Conclusion

This chapter 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 (previously described in Chapter 3) 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 exper-

ively 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.
ience improved, by understanding the tools and features that a community of users 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, which are 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 frustration coming 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 the intended audience.

To help to better align digital editions with 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 targeted at people who feel dissatisfied with existing digital editions. These and other possibilities speak to a field of research deserving of greater attention.
Figure 4.17: Answers recorded for survey question 17.
Figure 4.18: Answers recorded for survey question 18.
Figure 4.19: 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.
Chapter 5

Reassessing the oldest surviving manuscript of *De civitate Dei*: MS XXVIII(26), Biblioteca Capitolare di Verona

MS XXVIII(26), also known as Manuscript *Veronensis*, is the oldest surviving manuscript of Augustine’s *De civitate Dei* (The City of God).\(^1\) Dating back to the early Fifth Century AD, it is housed in the chapter library of Verona, Italy. Contemporary to Augustine himself, MS XXVIII(26) is an exceptionally exciting object of study. However, owing to a lack of evidence, scholars have long argued over its provenance. In a quest to understand the manuscript’s history, philologists and palaeographers have studied its script, language and genealogy. So far, no study has exploited digital technologies and resources to shed new light on the manuscript.

The purpose of this chapter is to reassess MS XXVIII(26) by bringing together all the accessible literature and to clarify why a digital edition of this manuscript is necessary. The information presented in this chapter was mostly gathered from

---

\(^1\)The Trismegistos reference number of this manuscript is TM 66599. Available at: http://www.trismegistos.org/ldab/text.php?tm=66599 (Accessed: 10 January 2014).
a variety of multilingual print publications scattered across different European libraries and institutions. Very little about this manuscript is, in fact, available online. This chapter looks at fourth and fifth century Verona, its place within the late Roman Empire and its relationship to the North African provinces. Next, it presents extant research, describes the manuscript itself and its current conservation status, explains why MS XXVIII(26) is a good candidate for a digital scholarly edition, and suggests avenues for future research which might help shed more light on the uncertain provenance of MS XXVIII(26).

Access to the manuscript was often difficult and inadequate. The short opening times of the library, combined with my life abroad and the manuscript’s frail condition, restricted access to the manuscript to very few visits. As a consequence, I purchased a CDROM from the Biblioteca Capitolare of Verona containing digital images of MSXXVIII(26), which were taken in 2003. However, the low quality of the images prolonged transcription efforts and was not sufficient to decipher those parts of the manuscript containing weathered text. Permission to re-image the manuscript at my own expense was denied and only partial reproduction of the 2003 digital images for the pilot digital edition was granted. More information about the implications of these limitations on the creation of a digital edition is provided in Chapter 6.

5.1 Verona in the 4th and 5th centuries AD

The imperial road to Italy goes from Munich across the Tyrol, through Innsbruck and Bozen to Verona, over the mountains. Here the great processions passed […] D. H. Lawrence²

In the fourth and fifth centuries AD, Verona, a pagan city of roughly 15,000 inhabitants,³ was an active trade and cultural centre. Built to resemble Rome,⁴

---
⁴Verona is sometimes referred to as Roma Minor in that it is considered to be the second city in Italy for Roman archaeological remains after Rome (Re, 2010, p. 74). Turrini reports Veronese
Verona was situated at the intersection of three arterial Roman roads: the Via Postumia, the Via Gallia and the Via Claudia Augusta, around which it grew and developed into an orthogonal grid (Magli, 2008, p. 66). Travellers and tradesmen would pass through the city on their way to the northern and southern Italian provinces, following these three important routes (see Figure 5.1). Its level land location, backed by hills and bathed by the river Adige, earned Verona its reputation and military prestige (Solinas, 1981, pp. 136-137, 151-152).

Verona also boasted one of the earliest, if not the oldest, writing centres in Europe (Biblioteca Capitolare di Verona, 1981, p. 13). Here, scribes curated classical, Christian and pagan literature, which renowned thinkers and teachers would peruse to advance their classical education (Graham, 2011, p. 154).

Migration to Europe was not uncommon for North African learned men in the fourth and fifth centuries. Whether prompted by political turmoil, or summoned to take on an important office, Zeno of Verona (300-371), Augustine of Hippo (354-430) and other African dignitaries travelled to Hispania, Gaul and Italy to spread the word of Christ through monasteries and apothecae (Graham, 2011, p. 152). Here, clerics, monks and scribes would copy the manuscripts they received from the Roman provinces, particularly from North Africa (Graham, 2011, p. 152).

[...] Italy, and southern Italy in particular, was one of the prime destinations for the dissemination of African texts to Europe in late antiquity and the Middle Ages. [...] The textual connection between Africa and Italy owed much of its significance to the seat of the Catholic Church at Rome, and the mutual culture of learned intellectuals and churchmen that bound Italy and Africa even into the

---

5The fall of the Roman Empire started with the vandal invasion and sack of Rome in 410. The Vandals are known to have entered North Africa in 429 (Merrills and Miles, 2010, pp. 52-53). For a complete history of the fall of the Roman Empire and of the Vandal occupation of Northern Africa see Heather (2006).

6For a detailed exploration of the Mediterranean migratory and commercial movements in the fourth and fifth centuries AD, see Wickham (2005).
seventh century. (Graham, 2011, p. 158)

5.2 The Biblioteca Capitolare

The *Schola sacerdotum sanctae Veronensis ecclesiae*, Verona’s *scriptorium* or calligraphic centre, is situated in the city centre next to the cathedral, the *Duomo*.\(^7\) Interestingly, the Duomo itself was erected next to the city’s first early Christian basilica, whose construction is attributed to Saint Zeno (300-371/372), the eighth

St Augustine and his *De civitate Dei*

Before examining the manuscript itself, some preliminary remarks about Saint Augustine and his *De civitate Dei* are necessary for the sake of orientation. Saint Augustine of Hippo was a theologian and philosopher. Referred to by Adamson (2013) as “a giant of the history of Christianity”, Augustine’s works are still widely read today and have had a great influence on the development of Western Christianity and philosophy (The Stanford Encyclopedia of Philosophy Online, 2010). Born a pagan in AD 354 in Thagaste,¹¹ at the age of twenty Augustine moved to Italy to teach rhetoric and oratory in Rome and Milan. Before returning to Africa in 388, he converted to Christianity and baptised his son

---

⁸Saint Zeno personifies Verona’s early and close relationship with the North African provinces.


¹⁰Uncial, Caroline and Humanistic, to name but a few.

¹¹Known as Souk Ahras in modern Algeria.
Adeodatus, who, however, died shortly after (AD 390). It was not until the late fourth century that Augustine firmly established himself within the wider Christian community: in AD 400 he was ordained bishop of Hippo Regius\textsuperscript{12}, where he had founded a monastery in AD 391 and where he produced his most influential writings, the \textit{Confessiones} and \textit{De civitate Dei}. Augustine died in AD 430 and was canonised in the twelfth century once the official canonisation process had been instituted by Pope Alexander III (Kemp, 1945).

Set against the backdrop of the fall of the Roman empire, \textit{De civitate Dei contra Paganos},\textsuperscript{13} commonly known as \textit{De civitate Dei}, is one of Saint Augustine’s major endeavours. Mainly crafted as a written attack against the religious beliefs and social values of Paganism and against those philosophers who supported it, \textit{De civitate Dei} examines the relationship and battle between the City of God (Christianity) and the City of Man (Paganism). The former is eternal and deals with the fundamental questions of virtue and vice, the latter is mortal and concerns itself with earthly needs (Adamson, 2013). But \textit{De civitate Dei} is not only a polemic. Augustine carves out some space to survey the history of mankind and of Rome, and to take the reader on a journey through post Pre-Socratic philosophy (Adamson, 2013). \textit{De civitate Dei} was written over a period of ten years sometime between AD 413 and 427 in Hippo Regius, and consisted of twenty-two books (see Appendix B).\textsuperscript{14} The publication of \textit{De civitate Dei} is documented by Augustine himself in a letter he wrote to an African priest named Firmus (Jones, 1965, p. 142). The letter (Hamman, 1960, columns 1373–1375) reveals Augustine’s wish to produce two types of edition for \textit{De civitate Dei}. Below is a reproduction of Jones’ summary of that letter:

\begin{footnotesize}
\begin{itemize}
\item \textsuperscript{12}Known as Annaba in modern Algeria, Hippo Regius was a major Roman African city. Augustine lived and died there (see Appendix B).
\item \textsuperscript{13}“The City of God against the Pagans”, in English
\item \textsuperscript{14}The number twenty-two is not accidental. \textit{De civitate Dei} is fraught with numerological references to, amongst others, God, the Trinity and the Bible; for a detailed exploration of Augustine’s use of numbers in \textit{De civitate Dei} see McCallin (1966).
\end{itemize}
\end{footnotesize}
Figure 5.2 implies two separate manuscript traditions, exemplified by two of the earliest manuscripts of De civitate Dei: Codex L(yon) 607 completely preserves books 1-5, suggesting adherence to the five-volume tradition;\textsuperscript{15} Codex C(orbie) 12214, on the other hand, completely preserves books 1-9, thus pointing to the two-volume tradition.\textsuperscript{16} As Jones explains, this dual format survives in later documents, thus corroborating the theory of a twofold manuscript tradition (Jones, 1965, p. 144).

When writing about Possidius’ Vita Augustini, Hamilton quotes and translates the following passage from Augustine’s good friend (Hamilton, 2004, p. 96):

> In this way, those who love God’s truth more than temporal riches will be able to choose the ones they wish to read and become acquainted with; then they may apply to the Church of Hippo to make a copy, or they may visit the library in Hippo, where they may find more correct copies, or may make inquiry wherever they can, and, having found what they want, may make a copy and keep it, and generously allow others to copy in turn. [Vita Augustini, 18.10]

\textsuperscript{15} Manuscript L 607 dates back to the sixth century and is today kept in the Bibliothèque Municipale de Lyon. A digitised version of the manuscript can be accessed here: \url{http://florus.bm-lyon.fr/visualisation.php?cote=MS0607&view=5} (Accessed: 25 April 2014).

\textsuperscript{16} Manuscript C lat. 12214, also known as the Corbie manuscript, also dates back to the sixth century. Part of it is housed in the Bibliothèque Nationale de France, part in the National Library of Russia (Leningrad Q.I.4, 43 ff.). A stable catalogue record can be viewed here: \url{www.trismegistos.org/text/66882} (Accessed: 25 April 2014).
While writing after Augustine’s death (see Appendix B), Possidius’ words describe Hippo Regius as a well-established library known for producing reliable editions of Augustine’s works (Gamble, 1995, p. 139). The monastery at Hippo “was the first one in North Africa and the parent of the other North African monasteries” (Weiskotten, 1919, p. 11). As was the practice with Augustinian works, Augustine is reported to have sent copies of *De civitate Dei* to trusted friends who would act as his ‘agents’ and allow copies to be taken (Gamble, 1995, p. 137). Legitimate editions were those whose copying had been authorised by the author himself and his agents (Gamble, 1995, p. 137). It is also very likely that Augustine deposited a copy of *De civitate Dei* in the library of Hippo in order to reach an even wider readership (Gamble, 1995, p. 136).

5.3.1 The tradition of *De civitate Dei*

The manuscript tradition of *De civitate Dei* spans a period of ten centuries and, while extensively discussed (Chadwick, 2004, p. 341), to this day only one *stemma codicum* (see Figure 5.3) or family tree of this work has been produced (Alexanderson, 1997). The author, Bernhard Dombart (1832-1907), dedicates the first thirty-four pages of his edition (Dombart, 1908) to examining this tradition but his stemma is based only on the first two books of *De civitate Dei*, thus excluding MS XXVIII(26), which preserves only books 11-16. Furthermore, as Alexanderson (1997) reports, Dombart’s model does not reflect correct and incorrect readings and cannot therefore be considered the most reliable of models. Nevertheless, Dombart’s stemma represents the first and only attempt to address a very elaborate study.

---

17 As previously mentioned, the archetype or first text of *De civitate Dei* was produced between 413-427 AD. The *editio princeps* of first edition was published in 1467 by Konrad Sweynheym and Arnold Pannartz. A digital copy of this volume is available here: https://archive.org/details/OEXV6RES (Accessed: 23 January 2014).

18 The stemma is available online via Google Books (s.v. *stemma codicum de civitate dei*) (Accessed: 20 January 2015).
5.4 MS XXVIII(26)

MS XXVIII(26) is the earliest surviving manuscript of *De civitate Dei*. Some scholars (see section 5.4.1) have said that it was written in Verona; recent scholarship points to North African origins; and other scholars still have suggested it originates in one of Augustine’s own scribal schools.

5.4.1 Previous research

Austrian philologist Emanuel Hoffmann (1825-1900) authored volume 40 of the *Corpus Scriptorum Ecclesiasticorum Latinorum* (CSEL). Written in 1899-1900, this edition of *De civitate Dei* predates the sporadic black retracing of the old lettering in MS XXVIII(26) (see section 5.4.2.4). In fact, CSEL’s critical apparatus docu-
ments features that are no longer visible today.\textsuperscript{19}

In the preface to his edition of \textit{De civitate Dei}, the German philologist Bernhard Dombart describes MS XXVIII(26) as belonging to the sixth century, while mentioning other scholars’ inferences: he reports August Reifferscheid, Wilhelm Weinberger and Emanuel Hoffmann dating the manuscript to as late as the seventh century. With regard to folia 1v-6r, Dombart believes these were written in the tenth century (Dombart, 1908, pp. 10-12).

In the 1930s, Elias Avery Lowe (1879-1969) openly recognised the fragility of his own conjectures. In the introduction to \textit{Codices Latini Antiquiores} (C.L.A.) IV he writes (Lowe, 1971):\textsuperscript{20}

\begin{quote}
I. If a manuscript has been preserved for generations in an ancient centre and local tradition connects it with that centre, the presumption is that it originated in that centre, provided, of course, there is no palaeographical evidence to the contrary. Our line of reasoning is: the manuscript is here, it has been here for centuries, it has no features which clearly suggest any other centre, hence we presume (even though we cannot prove it) that it was produced here. For example, when we find venerable Gospel books in such ancient ecclesiastical centres as Verona, Vercelli, Ancona, Brescia, and Perugia, we assume, in the absence of palaeographical and historical evidence to the contrary, that these volumes originated in the region which cherished them. If, in addition, entries made by later hands - they are mostly liturgical in content - point to the same region, the assumption gains considerably in probability. If, furthermore, the scripts of these Gospel books and of the added entries show noticeable differences from the same types of script in manuscripts known to have
\end{quote}

\textsuperscript{19}The bold lettering of the word \textit{LIQUOREM} in 17r.24, for instance, covers the original \textit{LIQUORIS}. See CSEL, volume 40, part 1, page 526: https://archive.org/stream/corpuscriptorum26wiengoog#page/n555/mode/2up (Accessed: 27 September 2014).

\textsuperscript{20}Two years after this text was written NUI Galway published a digitised version of Lowe’s descriptive entry for this particular manuscript. The page is available here: https://elmss.nuigalway.ie/catalogue/839 (Accessed: 17 August 2017).
originated, say, in France, Spain, or England, then we conclude that these Gospel books are of Italian origin, and probably local products. (C.L.A., Vol. 4, p. XII).

And when talking about MS XXVIII(26), Lowe proposes AD 420 as production year and makes the following statement:

Written doubtlessly in Italy, in a centre of high calligraphic standards […] The MS. is conspicuous for the excellence of its calligraphy rather than of its text. (C.L.A. 491)

Caroline Hammond Bammel (1944-1995) informs us that MS XXVIII(26) was copied from a previous exemplar. Her claims, while not supported by concrete evidence, fit with Lowe’s dating suggestion in that she too believes that the manuscript was produced shortly after 420 (Bammel, 1979, p. 441). Bammel also reports Lowe listing MS XXVIII(26) as North African, a fact she has not yet been able to corroborate herself (Bammel, 1979, p. 435). While not openly expressing her beliefs, Bammel acknowledges that evidence for African provenance is strong.

With his Survey of the Oldest Manuscripts of St Augustine’s De civitate Dei, Michael M. Gorman lists MS XXVIII(26) as North African. However, he does not provide his own reasoning behind this assertion.

Graham (2011) claims many manuscripts travelled from Northern Africa to Italy along established Roman trade routes, which passed through Naples and Rome.

People carried books and manuscripts along the routes of trade and pilgrimage, along channels of communication among scholarly and

---

21 This statement is corroborated in an earlier article by the palaeographer, who based his provenance analysis of various manuscripts on three characteristics of ‘Africanness’ put forward by Alexander Souter, a contemporary biblical scholar. While MS XXVIII(26) meets one of these requirements, Lowe concludes:

No one will venture to claim these MSS for Africa on the strength of this feature (capitals at the beginning of each page or each column). Some of them obviously originated in Italy, with which country their whole past history is bound up. Others are so like them as to justify their being ascribed to Italy. (Lowe, 1922, p. 404)
religious communities, and along the channels of imperial, papal, and ecclesiastical administration. (Graham, 2011, p. 153)

Furthermore:

[...] every North African text that survives in a European manuscript left Africa before the Arab conquest of the late seventh century. (Graham, 2011, p. 151)

While we have no clear indication as to the exact itineraries (Rouse and McNelis, 2000, p. 230), current scholarship in the fields of ancient geography, commerce and transportation allows us to map extant literary evidence and supports Graham’s argumentation, as exemplified by Figure 5.4 below. Indeed, Hammond Bammel informs that the geographical distribution of Augustinian manuscripts follows the North Africa, Campania, Rome and North Italy axis (Bammel, 1979, p. 441).

To sum up, dating attempts thus far have tended to place the manuscript in the second, rather than the first half of the fifth century. Again, although most of these conjectures seem to align, the absence of definitive proof opens them up for debate.

Cameron (2004) describes the difficulties in dating late-antique manuscripts, and MS XXVIII(26) is not the only manuscript Lowe had trouble placing (Cameron, 2004, p. 519). An early date cannot be excluded but cannot yet be proven either.

5.4.2 The manuscript

In his manuscript catalogue Codices Latini Antiquiores, the renowned palaeographer Lowe writes:

491.-Verona, XXVIII (26). Augustinus de civitate Dei (LIBB. XI-XVI).

UNCIAL SAEC. V1 (post AD 420)

22 ‘People’ included travelling monks, who were often given money to buy books (Lyons, 2011, p. 38).
23 Campania is a southern region of Italy. Its regional seat is Naples.
24 For a brief overview of palaeographical dating methods and modern dating techniques, see
Figure 5.4: Screenshot of an ORBIS map displaying the fastest (red), cheapest (green) and shortest (purple) summer routes from Hippo Regius to Verona. ORBIS is available at: http://orbis.stanford.edu/ (Accessed: 17 October 2014).

Fol. 244 (in the whole MS. foll. 253: foll. I-6v, containing the Capitula; fol. 102 and 248 are slips with Veronese minuscule saec. IX; foll. 252-3 is a bifolium containing on fol. 252v a prayer supposed to be in the hands of Ratherius who became bishop of Verona in 932); ca. 292x190mm. (195-205x120-130mm.) in 30 long lines. Ruling on the flesh-side, apparently 2 bifolia at a time, before folding. Single bounding lines enclose the text. Prickings to guide ruling run through the text—a practice of very ancient MSS. Gatherings of eights, with flesh-side outside, signed by a tiny Φ and a Roman nu-
meral in the right-hand lower corner of the last page. Running title, in small uncial, on openings between quires only (except at the end of quire XIII, where it is in half-uncial). Dignified colophons in large uncial in alternate lines of black and red between rows of bird-like flourishes. No punctuation, apart from an occasional medial point; mostly the space of one or two letters is left blank to denote a pause. Citations are numerous but unmarked (the flourishes in the margin of fol. 186v are added). Omissions supplied interlinearly. Abbreviations restricted to B., Q. for ‘bus’, ‘que’, FF for ‘fratres’ and to Nomina Sacra usually followed by a point: DNS, DNM (normally) = dominus, -um, but the older syllabic suspension DMN occurs for ‘dominum’ on fol. 121; SCIS = sanctis; XPS = christus; IS (for the normal IHS) occurs on fol. 192. Omitted M and N, at line ends, marked by a simple line after the vowel. Each page begins with a larger letter, even in mid-word. Red is used for the opening 2 lines of each book. No decoration. Parchment well prepared as for a MS. de luxe, but membranes vary in thickness, some fine, some quite stout. Ink, rich brown, scaled off on the flesh-sides and retracted. Script is broad and very expert uncial of the oldest type, recalling the Puteanus (Paris Lat. 5730) and the Lateranensis of Livy (C.L.A., I, No. 57). Here and there are marginal entries in contemporary half-uncial (fol. 31v) and cursive quarter-uncial, by very expert hands (see plate). The eighth-century added notes on foll. 143v and 186v seem Veronese. Numerous marginalia by Pacificus of Verona (see plate). (C.L.A., IV: 491)

The different quality of the parchment used suggests that the scribe used whatever material was available at the time.
5.4.2.1 Missing books

Of the twenty-two volumes of *De civitate Dei*, MS XXVIII(26) contains only 11-16. Where and when did the others go missing? The table below places MS XXVIII(26) (in red) within the two publication formats envisaged by Augustine, previously illustrated in Figure 5.2:

**Figure 5.5:** From top to bottom. The top row represents the two-volume format envisaged by Augustine for *De civitate Dei*. Here, volume one contains ten books and volume two contains twelve books; in this two-volume format, MS XXVIII(26), depicted in red, would be enclosed in volume two. Conversely, the lowest row of the table represents the five-volume format of *De civitate Dei* as envisaged by Augustine; here, MS XXVIII(26) would enclose volume three and the first two books of volume four. Table my own creation.

Unlike Hammond Bammel, I believe books 11-16 of MS XXVIII(26) to be an incomplete second volume of Augustine’s two-piece edition rather than a merge of volumes three and four of the five-piece edition. A merge, in fact, would not only affect the flow of the narrative but would also imply the presence of an incipit to book 17, which, however, is not present.

This conjecture, however, does not solve two further questions: what happened to books 17-22 (the remaining books from the second volume of the two-piece

---

25 The Biblioteca Capitolare houses another thirty-seven manuscripts containing Augustine’s works. These date between the sixth and sixteenth centuries (Adami, nd).

26 It was because of illegal copying of his *De Trinitate* that Augustine decided to publish *De civitate Dei* in instalments. He believed that such a division would not only allow to better monitor circulation but would also help readers split this long work into manageable chunks (Gamble, 1995, pp. 133-134).
edition) and 1-10? With regard to the former question, it is possible that only books 11-16 had been completed by the time Augustine wrote to Firmus. This might explain why an expected incipit to book 17 is not present at the end of book 16 in MS XXVIII(26) (Bammel, 1979, p. 435). Nevertheless, it seems unlikely that scribes would copy an incomplete work, if MS XXVIII(26) is in fact a copy. On the other hand, evidence also suggests that Augustine did not publish De civitate Dei as a complete work but in instalments (Green, 1959, p. 193). This hypothesis, corroborated by Gamble (Gamble, 1995, p. 135-136), is further confirmed by the Epistola ad Firmum itself, which was written in AD 426 (Lambot, 1939, p. 109), the year in which Augustine is known to have been producing books 11-22 of De civitate Dei (see Appendix B). If we assume a North African provenance for MS XXVIII(26), it is possible that MS XXVIII(26) had been broken down into smaller sets for travelling purposes and that only books 11-16 reached Verona:

Manuscripts travelled in books —the codex was an important tool of Christian culture— and as single quires or folios, and were always, of necessity, accompanied by people. (Graham, 2011, p. 153)

And:

It has been inferred […] that, at some time around the year 440, the entire collection of Augustine’s works may have been transferred to Rome from Hippo […]. (Vessey, 2015, p. 435)

There is no evidence at this time suggesting such a division took place with MS XXVIII(26) but, again, this possibility would perhaps justify the absence of an incipit to book 17, and suggest that 17-22 either went missing or survive in some unknown location.

As for books 1-10, the fact that book 11 in MS XXVIII(26) does not contain any introductory text or title page suggests that it is a continuation from previous volumes and that books 1-10 had already —logically— been written. As Augustine envisaged book 11 as the shared partition between the two- and the five-volume structure of De civitate Dei, we might accordingly expect some form of
introduction. Its absence leads us to conclude that books 1-10 and whatever introductory note might have preceded folio 7 also went missing.

5.4.2.2 Parchment

A close observation of the parchment reveals a number of holes. These could be stretched-out insect bites or accidental knife holes resulting from skin scraping (Avrin, 2010, p. 213). One can also spot vein furrows and follicles on the skin-side. Typically, pagination paired skin and flesh sides, whereby two skin pages were followed by two flesh pages (Ryder, 1960, p. 394). The pagination of MS XXVIII(26) follows this arrangement. The skin most certainly belonged to an ovine, with the arrangement of the follicles (in groups of three) pointing to goat-skin (see Figure 5.6).

The finest parchment, as Ryder informs, was derived from calves, not ovines. In fact, goatskin is coarser and thinner than that of a young bovine.

Parchment prickings were typically carried out in the margin. In MS

---

27 Stretching (“drying under tension”) was part of the parchment manufacturing process, where the parchment maker would dry the animal skin under tension by stretching it out on a wooden frame: “The effect of the stretching is to cause the collagen fibres of the skin to lie in sheets parallel to the skin surface. […] This explains why parchment […] can be torn with the hand into a number of thinner sheets” (Ryder, 1960, p. 394).

28 Veining is a common imperfection of parchment (Johnson, 2008). Parchment, or animal skin, has two sides: skin and flesh. The flesh side tends to be smoother and paler, whereas the skin side is more yellow in colour and can easily be identified by hair follicle marks (Avrin, 2010, p. 213).

29 This pairing system is called Gregory’s Rule, after theologian Caspar René Gregory, who was the first to discover this practice. For more information about Caspar René Gregory, see: http://www.jstor.org/stable/3141854?seq=1#page_scan_tab_contents (Accessed: 10 July 2015).

30 Any inconsistencies in this page ordering are likely traceable to the Vatican restoration (see section 5.4.3).

31 Veal skin has evenly distributed follicles; groups of three are distinctive of goats; sheep follicles are usually arranged in a circle; hair is also different between animals: goat’s is more angular while sheep’s more perpendicular (Ryder, 1960, p. 397). Sheep and goatskin were the most common source of parchment in continental Europe (Ryder, 1960, p. 394).

32 The terms parchment and vellum are often used interchangeably. The latter comes from the French word for calf, ‘vel’. For a detailed analysis on the difference between the terms parchment and vellum, see (Ryder, 1960, p. 393).

33 Pricking the parchment along the long edge of the page served to guide the subsequent ruling (see following footnote). Prickings were often, but not always (as in this case), cut out during the binding process. For more detailed information about prickings, see (Lowe, 1971, IV, p. VII).
XXVIII(26), however, prickings appear within the text itself, a decision perhaps dictated by a scribal desire to keep them hidden. In addition, contrary to common practice, the text is placed above the ruling (see Figure 5.7), not beneath.\textsuperscript{34}

In its complete form, MS XXVIII(26) would have totalled some eight-hundred folia.\textsuperscript{35} As one sheep or goat provided enough skin to produce four folios, the total number of goats needed to produce this manuscript would have been roughly two-hundred (Wallace, 2011).

\textsuperscript{34}Ruling was used to evenly place the text on the page and to help the scribe write in a straight line.

\textsuperscript{35}Author’s estimate based on the number of folia covering books 11-16.
5.4. MS XXVIII(26)

5.4.2.3 Text (books 11-16)

Figure 5.8: The front matter of MS XXVIII(26) reads: Beatus Augustinus de civitate dei libros VI ut (et) XI, XII, XIII, XIV, XV, XVI

Folia 1v-6r. The introductory folia 1v (verso) to 6r (recto) were added in the eighth/ninth centuries and contain the capitula or summaries of books 11-14. The script, produced by a single hand, is Veronese minuscule (Venturini, 1929, p. 149), a variation of Caroline minuscule, thus implying that the manuscript was in Verona by the eighth century (Bammel, 1979, p. 443). Caroline minuscule became widely adopted in the ninth century in all scriptoria in France, Germany, and in northern and central Italy, including Verona, where the script was influenced by, amongst others, Irish insular script, a minuscule high class script whose presence in MS XXVIII(26) is noticeable in the use of marginal three-dot markers (∴ or ∵) serving as decoration or as clause endings (Bischoff, 1990, p. 117). Some of the features common to this script are (as also found in MS XXVIII(26):

Figure 5.9: AE diphthong represented by a minuscule letter E with descender, also known as e caudata (folio 1v, line 7).

---

36 Verso is Latin for ‘back’ and recto for ‘front’. The capitula are fragmentary, in that they do not contain capitula for books 15 and 16. Folia 102r and 248b also contain capitula albeit in fragmentary or partial form.
Figure 5.10: Carolingian and insular minuscule letter G (folio 1v, line 28).

Figure 5.11: The ET ligature looks like an ampersand (folio 1r, line 25). The symbol we use today as a replacement for the English conjunction AND is in fact a combination of the uncial characters e and t, which make up the Latin conjunction et (= and). In MS XXVIII(26) this ampersand symbol is also inconsistently used to replace -et verb endings or any other occurrence of this letter pairing.

Figure 5.12: The minuscule letter r is combined with an especially long i (folio 2r, line 4).

Moreover, the letter I is often capitalised; its use in mid-word and also at the beginning of a word is not consistent, with the exception of the pronoun eius which always presents a capitalised I. The annotations and decorative features covering the manuscript also belong to the ninth century. MS XXVIII(26) contains a total of 248 scholia or annotations (see section 5.4.2.4) as well as critical signs used by multiple scribes and scholars as reference points to help readers navigate the text (Venturini, 1929, p. 150) (Carlquist, 2004, pp. 107-108). The Veronese script occasionally employs decorative features as both macra and reference markers, as illustrated in Figures 5.13 and 5.14 (Venturini, 1929, pp. 79-80):\textsuperscript{37}

\textsuperscript{37} Appearing above a word, a macron (plural macra) is a horizontal marker or line used to flag an abbreviation (more specifically, suspensions and elisions).
5.4. MS XXVIII(26)

Figure 5.13: Two types of decorative markers, both labelling the beginning of a new chapter: on the left (folio 2v) the marker appears to be in the shape of a leaf; on the right (folio 12v), the capital letter C and the Roman numeral seven indicate Capitulum seven starts with the word UIDEMUS.

Figure 5.14: The title on folio 1v contains decorative motifs acting as macra. LIB is an abbreviation of LIBRI and DI of DEI.

5.4.2.4 Scholia

During the ninth century Verona produced 218 codices under the directorate of Archdeacon Irenaeus Pacificus (776-844). Formerly a bishop, Pacificus became archdeacon in 801 (Turrini, 1948, p. 59). A pupil of Egin of Reichenau, Pacificus authored, corrected and annotated manuscripts in a style that very much resembled that of the German monastery (Venturini, 1929, p. 30). As previously stated, MS XXVIII(26) had been in the Veronese library before his arrival and was one of the Archdeacon’s objects of study, as attested by the scholia ad-

---

38 His epitaph, inscribed on the wall of Verona’s Cathedral, reads: ‘[…] Nullus unquam sic peritus in tantis op(eratoribus) / Bis centenos terq(ue) senos codices(que) fecerat’ (Venturini, 1929, p. 2).

39 St Egin, thought by some to be a relative of Charle Magne, was bishop of Verona from 796 to 799 (Hlawitschka, 1993). The strong relationship of Verona with its sister monasteries of Reichenau and St Gallen is confirmed by Egin himself who carried a number of books back with him to Reichenau after his resignation from episcopal duties in 799 (Rouse and McNelis, 2000, pp. 230-231). Egin was succeeded by Ratoldus (a Benedictine monk, also from Reichenau) who was bishop until 840. Ratoldus is known to have been in close personal and professional relations with Pacificus (Venturini, 1929, p. 71).

40 For example, the Glossae Super Exodum contained in Codex LXIX (66) of the Biblioteca Capitolare. A plate of this manuscript can be viewed in (Spagnolo, 1996, p. 129). Insular script developed in Ireland and England between the 7th-9th centuries. It was well known in the monastic scriptoria of Bobbio, Luxeuil, St Gallen, Corbie, Reichenau, Fulda, and Regensburg (De Lasala, 2010, p. 58). Verona’s working relationship with these monasteries, particularly Bobbio, serves to explain the script’s influence on ninth century Veronese scribes.
ded by both Pacificus and other hands (Venturini, 1929, p. 68). These are often preceded by an R, an abbreviation of the word require (= check), and vary in content (mythological, philosophical, philological, moral, legal, critical, scientific and theological) and in length. The handwriting of folia 1v–6r suggest these were also written by Pacificus. As Venturini writes (Venturini, 1929, p. 145), navigation and reading tools such as on these pages could only have been composed by someone familiar with the text. Yet, the spelling errors, abbreviation inconsistencies, missing capitula and incorrect numbering would speak to distraction. Despite the striking calligraphic resemblance, the number of what seem to be distraction errors suggests these pages were perhaps written under (Pacificus’) authorial dictation.

**Folia 7r-251v.** The ancient text contained in these pages is written in *scriptura continua* or, in other words, as a long string of characters with no punctuation. Some punctuation marks, namely semicolons, were added to MS XXVIII(26) by a later scribe as a reading aid. Reading *scriptura continua*, as Gamble reveals, was highly interpretative insomuch as “[the] text […] presented a greater range of interpretive options and demanded more hermeneutical decisions of the reader than modern texts do”. One way of reading continuous script was to sound the syllables, leading one to argue that MS XXVIII(26) was intended to

---

41 Including perhaps that of Ratherius, the tenth century bishop of Verona (Turrini, 1948, p. 19). Codex Paris Lat. 1924, f. 1 contains an authentic Pacificus signature. The handwriting is identical to that of the many scholia present in MS XXVIII(26) and leads us to conclude that it was indeed Pacificus who authored the ninth century annotations. An image of this snippet of text can be viewed in Turrini (Turrini, 1948, p. 87).

42 Such as *axtra* in folio 1v, line 14 instead of *astra*.

43 *Capitulum* XI of book 11, for example.

44 Book 11 contains two capitula VIII.

45 Gamble writes: “Both Augustine and Jerome relied heavily on stenographers and copyists. For Augustine in particular, composing was always a matter of dictating (*dictare*), and once a fair copy was transcribed from stenographic notes, of correcting (*emendare*). Once transcribed and corrected, texts were *emendatoria exemplaria* (= corrected exemplars), and it was from these that copies were made. As bishop of Hippo, Augustine was well furnished with stenographers and copyists.” (Gamble, 1995, p. 139) The term *scribe* comes from the Latin *scribere*, a verb used to describe the act of writing (Fleischman, 1990, p. 20).

46 The location of a word-separating punctuation mark is indicative of the length of the pause: when at the top, the dot signifies a long pause; in the middle, medium long pause; at the bottom, short pause (De Lasala, 2010, p. 42).
be read out loud (Gamble, 1995, p. 204).\textsuperscript{47} The text is written in uncial, a script which was first developed in North Africa in the fourth century AD (Petrucci, 1989, p. 63).\textsuperscript{48} Uncial is a majuscule script known for its characteristic rounded letters, particularly the letters $A$, $D$, $E$ and $M$ (see Figure 5.15):

\textbf{Figure 5.15:} Characteristic uncial letters $D$, $E$, $A$ and $M$ in line 13, folio 73v.

With the exception of the minuscule letters $h$, $p$ and $q$ (see Figure 5.16), the majuscle letters run within two parallel lines or in what is known as a \textit{bilinear} scheme.\textsuperscript{49} It is perhaps this scheme that forces the roundness of the lettering.

\textbf{Figure 5.16:} Minuscule letters $Q$ and $H$ in line 14, folio 57r.

Uncial was extensively used in the scriptorium of Verona during the fifth century (Piazzi et al., 1986, p. 36). Indeed, Italian and African scriptoria were the major producers of uncial codices during the late antique period (Petrucci, 1989, p. 64). Each page of MS XXVIII(26) begins with a capital letter, irrespective of the word or sentence’s position within the text (Lowe, 1925, p. 197). This form of capitalisation is a practice of great antiquity and is visible in numerous fifth century uncial manuscripts (Lowe, 1925).

The original ink, gall ink, is dark brown.\textsuperscript{50} Some folia present darker, almost

---

\textsuperscript{47}Ehrman writes: “For in the ancient world ‘reading’ a book did not mean, usually, reading it to oneself; it meant reading it aloud, to others. One could be said to have \textit{read} a book when in fact one had \textit{heard} it read by others. […] books were almost always read aloud in social settings, such as in settings of worship.” (Ehrman, 2005, pp. 41-42).

\textsuperscript{48}Uncial became the main Christian script and was adopted in the entire Latin West from the fourth until the ninth-tenth centuries (Petrucci, 1989, p. 64).

\textsuperscript{49}When the letters are made to fit between two parallel lines. \textit{Minuscule} script (such as half-uncial), with its ascenders and descenders, stretches across four parallel lines.

\textsuperscript{50}Gall-ink was made out of crushed galls (abnormal growths found on plants), which contain tannic acid. This acid, mixed with gum arabic, water and iron sulphate, produces gall ink. This type of ink was the most commonly used writing ink of the Middle Ages (Lyons, 2011, p. 43).
black strokes. This is the result of later, rough and irregular retracing, possibly dating back to the 1920s when the manuscript underwent restoration at the Vatican Library (see section 5.4.3). This second layer of acidic iron ink, combined with the original gall ink, triggered a chemical reaction resulting in parchment corrosion (see Figure 5.17).

**Figure 5.17:** Ink corrosion on folio 74r (lines 27-30). The different inks are clearly visible: the original gall-ink is visible at the beginning of sentence 30 with the first half of the word *corporales* (*CORP-*) ; the second half of the word (*ORALES*) and, indeed, the rest of the sentence was retraced and presents noticeable signs of corrosion.

Corrosion is occasionally accompanied by heavy moulding, so much so that in some cases the text and the scholia are almost completely obliterated (see Figure 5.18). This is where high-resolution imaging, and multispectral techniques in particular, could help restore the text.

**Figure 5.18:** Part of the text in folio 61r is completely hidden by mould. Even the scholia (blue focus box) are barely visible.

Crimson mould is the result of water stains or humidity, with darker areas be-

---

51 Multispectral imaging is an image-capture technique that exploits electromagnetic wavelengths and frequencies, including infrared and ultraviolet, in order to facilitate the extraction of information that the human eye cannot identify. For a more detailed explanation of multispectral imaging techniques, see Terras (2012b).
ing indicative of prolonged exposure to the liquid (Polacheck et al., 1988, p. 93). Where particularly acidic, the iron gall ink repels the mould causing what book conservators call the 'haloing' effect. The abbreviated word INCIPI'T in Figure 5.19 is a good example of haloing.

![Figure 5.19: The halo surrounding the abbreviated word INCIPI'T on folio 76v is caused by the reaction of the mould against the iron gall ink.](image)

With regard to the Latin, this manuscript makes no distinction between the letters U and V. Up until the sixth century, these two letters were pronounced the same but their position within a word was fixed: U appeared in the middle of a word, V at the beginning. Similarly, the letters I and J sound identical but occupy different positions within a word; the former appears only at the beginning, the latter in the middle (De Lasala, 2010, p. 40). The text contains rare abbreviations such as is for Iesus and dnm for dominum, as well as standard abbreviations like the middle dot used to replace a -que or -bus, and the consonant suppression (= notae communes) at the end of a word (final M and N is replaced by a top line or macron) (Bammel, 1979, pp. 430-431). From the fourth century AD onwards, Christian texts also abbreviate sacred names, known as the nomina sacra: MS XXVIII(26) contains many of these (see Table 5.1), including the Greek XPS (= Christus), which is distinctive of early Christian works (Brown, 1990, p. 5). Not all of them, contrary to Bammel’s observation (Bammel, 1979, p. 462), are followed by a point in MS XXVIII(26).

---

53 MS XXVIII(26) does not make a distinction between omitted m and n. These can only be gleaned from the knowledge of the language. For an analysis of the distinction between omitted m and n, see Lowe (1925).
54 For a detailed account of fifth century nomina sacra, see Bammel (1979).
55 An example can be found in folio 42r, line 20.
Table 5.1: List of abbreviations occurring in MS XXVIII(26) with the corresponding expansions.

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Expansion</th>
</tr>
</thead>
<tbody>
<tr>
<td>xpm</td>
<td>ihm; is</td>
</tr>
<tr>
<td></td>
<td>Christum Iesum; Iesus</td>
</tr>
<tr>
<td>do; di; dm; ds</td>
<td>deo; dei; deum; deus</td>
</tr>
<tr>
<td>sps; spu</td>
<td>spiritus; spiritu</td>
</tr>
<tr>
<td>sci; scs</td>
<td>sancti; sanctus</td>
</tr>
</tbody>
</table>

The presence of ligatures is also notable, particularly at the end of a line; the final -NT in words like PUTANT or CREDANT, the ligature -US in words like UNUS, and -UT in words like SICUT. Decorations also include Christograms, that is, abbreviations for the name of Jesus Christ (see Figure 5.20).

Figure 5.20: This Christogram, drawn on folio 7r, appears to be an original feature (fifth century).

Folio 252v. Folio 252v was likely written in the eighth century by one of the hands that annotated the uncial text. The script, pre-Caroline German minuscule, is often found in manuscripts attributed to Reichenau and thus bears witness to the already mentioned strong relationship between Verona and its German sister scriptorium (Brown, 1990, p. 44). This folio contains a variation on an oration of Catholic liturgy.  

---

56 This oration can also be found in Moeller, E. E., Clément, J. M., Coppierst Wallant, B. (eds.) Corpus Orationum, n. 4227.
Figure 5.21: Folio 252v dates back to the eighth century and contains a Catholic liturgy.
5.4.3 Conservation status

As per the conservation report (see Appendix C), the manuscript underwent restoration at the Vatican Library in the 1920s. Conservation practices at the time caused irreversible damage to the manuscript; the acid used to retrace the text to improve readability when combined with the ink was highly effective but provoked a chemical reaction causing the text and its physical support to gelatinise. In addition, as previously discussed, some folia have been greatly affected by mould, a common result of humidity.

In early 2012 I contacted the Vatican Library with the intent of obtaining more information about the restoration work and the specific measures taken. A swift reply from Dr Angela Nuñez Gaitan revealed the absence of any documentation with the exception of a small excerpt, which reads (Gaitan, 2012):57

\begin{quote}
\end{quote}


The term ‘rebound’ in this short report confirms that the Vatican removed the manuscript binding prior to restoration and re-attached it once the restoration had been completed. Of course, damage was also probably caused by the numerous disasters that befell the Verona library between the seventeenth and twentieth centuries: misplacement in 1625, flooding in 1882 and war bombings in 1945.59 These ruinous events, combined with the manuscript’s age and natural

---

57 At the time of writing, Nuñez Gaitan is Head of the Conservation Workshop at the Vatican Library.
58 Translation by the author.
59 As Del Re recounts (Re, 2010, p. 75), Monsignor Agostino Rezzani, who was the librarian at the time, hid a number of the Library’s codices and manuscripts as a precaution against the Library’s relocation to a nearby building, which represented a security hazard and could have
decay, bind MS XXVIII(26) to the archive of the Biblioteca Capitolare of Verona.

So what has the discussion thus far revealed about the provenance of MS XXVIII(26)? The historical and palaeographical evidence is not conclusive, and most experts prefer to place its production in Italy, although North African origins cannot be excluded. In the next section, MS XXVIII(26) is compared against contemporary exemplars in an effort to shed more light on its provenance. However, the number of pre-800 AD manuscripts is very small, so any resemblance should be carefully evaluated.

### 5.4.4 Comparing MS XXVIII(26) to coeval manuscripts

Let us compare MS XXVIII(26) against manuscripts whose analogous date and provenance have been proven. As De Lasala explains (De Lasala, 2010, p. 18):

Addirittura, le più antiche copie che abbiamo della maggior parte delle opere di Sant’Agostino sono state probabilmente elaborate in terra africana. [È stata riconosciuta una di queste copie in un codice in scrittura onciale di San Pietroburgo —Russia— contenente le quattro opere che compose Agostino d’Ippona durante il primo anno del suo episcopato (395-396). Uno degli scrittori di quel codice ha adoperato un tipo di onciale “africana”, come quello usato nei codici di San Cipriano, il che conferma il carattere africano originario del codice].

The following is my English translation of De Lasala: “The oldest copies we possess of the majority of Augustine’s works were probably produced on African soil. [One of these copies was identified in an uncial codex belonging to St Petersburg —Russia— containing four works Augustine of Hippo wrote during the first year of his episcopacy (395-396). One of the authors of that codex adopted the type of ‘African’ uncial that was also used in the codices of St Cyprian, thus confirming the native African character of the codex].”

---

60 The following is my English translation of De Lasala: “The oldest copies we possess of the majority of Augustine’s works were probably produced on African soil. [One of these copies was identified in an uncial codex belonging to St Petersburg —Russia— containing four works Augustine of Hippo wrote during the first year of his episcopacy (395-396). One of the authors of that codex adopted the type of ‘African’ uncial that was also used in the codices of St Cyprian, thus confirming the native African character of the codex].”
Table 5.2 below lists manuscripts contemporary to MS XXVIII(26) that I was able to find in all the accessible literature:\(^{61}\)

**Table 5.2: Manuscripts contemporary to MS XXVIII(26).**

<table>
<thead>
<tr>
<th>Manuscript</th>
<th>Date</th>
<th>Provenance</th>
<th>URL</th>
</tr>
</thead>
</table>

Now in the National Library of Turin, *Codex Bobbiensis* takes its name from the monastery in the Italian town of Bobbio. Bobbiensis contains the Latin text of the Bible and, as is often the case in manuscript studies, its birth date has been a topic of discussion for many years, with academics arguing over the fourth and fifth centuries (Cipolla, 907a). The Italian scholar Cipolla attributes Bobbiensis

---

\(^{61}\)I travelled to the Warburg Institute, London, in June 2014 to specifically access all available literature and facsimiles pertaining to this section. Multiple trips to view each manuscript listed in the table were not possible owing to the considerable financial resources required, not least the procurement of all necessary access permissions.
to the fifth century inasmuch as, just as with MS XXVIII(26), it does not present any of the elegance typical of a fourth century manuscript (Cipolla, 907a, p. 70). The script has been incontestably identified as North African due to provincial and curious symbols used in the *nomina sacra* and to the uncial characters, which differ from the Italian uncial (Traube, 1907, p. 138 ff.).
As can be seen from the images above, African uncial characters (see Figure 5.22)
are narrower and more angular than Italian uncial (see Figure 5.23), which shows rounded and fatter letters. The letter \( d \) in both scripts features a sharp bow; both also contain serifs, which, however, present themselves more prominently in Italian uncial with the letters \( f, l, s \) and \( t \).

![Figure 5.24: MS XXVIII(26), extract from folio 57v.](image)

Upon comparing MS XXVIII(26) (see Figure 5.24) to these manuscripts, one notices a mix of the two scripts; like Italian uncial, the characters in MS XXVIII(26) are tidier and more rounded but do not present the Italian characteristic serifs; the letter \( e \) is clearly African, as is the overall appearance, with descenders and ascenders often leaping over the bilinear scheme. Codex Bobbiensis and MS XXVIII(26) also share the \( NT \) ligature but not the positioning of the \( m \) and \( n \) suspension markers. Italian uncial, in fact, tends to place them after the preceding vowel, whereas African places them on top of the vowel. It is here that MS XXVIII(26) deviates from an African style insomuch as it follows the Italian suspension scheme. Nevertheless, while presenting a mix of features, and on the basis of this comparison, MS XXVIII(26) would appear to be more African.
The *Bologna codex* contains the works of African Christian author Lactantius. This codex dates back to the second half of the fifth century and showcases a more elaborate version of African uncial (Lowe, 1938, p. 3). With the exception of the long and curly descenders and the tighter bilinear scheme, the script of Bologna 701 is very similar to that of MS XXVIII(26).

The *Leningrad Codex* offers the best evidence in support of MS XXVIII(26)'s African provenance. One of the earliest manuscripts we know of that was sent from Hippo Regius to Italy (Naples), its script is identical to that of MS XXVIII(26) (Graham, 2011, p. 154). The striking resemblance places MS XXVIII(26) not only in North Africa, but in Hippo Regius, thus drawing nearer to Augustine himself (Lowe, 1966, p. 8). To the best of my knowledge, no scholar has attempted to prove a Hippo origin for MS XXVIII(26).

Thirty-three folia of a codex preserving the letters of the African author Cyprian are spread across three different libraries in Turin, Milan and the Vatican (Lowe, 1947, pp. 16-17). Lowe dates this codex to the fourth-fifth centuries. Its thick-stroke writing style resembles that of Bobbiensis, where the uncial script appears narrower and more elongated than that of MS XXVIII(26) and Leningrad. As reported by Lowe, the palaeographical similarity to Bobbiensis would favour the African side of the provenance debate (the opposing view being, once more, Italy) but this remains a conjecture (Lowe, 1947, pp. 16-17).

Another exemplar dating to the fifth century whose provenance remains uncertain is *Codex Palatinus*. As per the British Library record of one of its leaves (Add MS 40107, folia 1r and 1v), Palatinus contains the Old Latin text of the Gospels composed in silver lettering on purple vellum. The uncial characters are less rounded than the previous codices, and present strong, thick strokes. The rigour with which the scribe inscribed the text onto the vellum is characteristic of, so-called, purple codices, high-class objects intended for the wealthy. The similarity

---

62 Lowe writes: "Foll. 33 survive: 24 at Turin, 8 at Milan and one in the Vatican […]".
to MS XXVIII(26) here is visible but not obvious. Of all the exemplars more or less contemporary to MS XXVIII(26), Codex Palatinus appears to be the least similar.

Finally, the digitised Codex Sangallensis 1395 contains fragments of the oldest Vulgate version of the Gospels, and is currently catalogued as produced in northern Italy (Verona?) in about 410/420 (see 5.25).64 Upon comparing the script of this codex to that of MS XXVIII(26) one notices many differences. The Carolingian G, the long S and the generally more minuscule script in Codex Sangallensis 1395 (visible in the letters A and U, for instance) do not appear in MS XXVIII(26). This palaeographical evidence does not speak in favour of a shared scriptorium.

Had all the above manuscripts been copies of De civitate Dei, a philological examination of the scribal errors contained therein would have perhaps helped us

---

64I came across the manuscript on 30th June 2015 thanks to a discussion on Twitter: https://twitter.com/EdvanderVlist/status/615796892430012416?ref_src=twsrc%5Etfw. The e-codices manuscript record, however, informs that the images have been online since 31st July 2009. The record is accessible here: http://www.e-codices.unifr.ch/en/searchresult/list/one/csg/1395 (Accessed 25 July 2015).
better identify shared provenance. In fact:

[...] during the early centuries of the church, Christian texts were copied in whatever location they were written or taken to. Since texts were copied locally, it is no surprise that different localities developed different kinds of textual tradition. That is to say, the manuscripts in Rome had many of the same errors, because they were for the most part “in-house” documents, copied from one another [...]. (Ehrman, 2005, p. 72)

However, as they all carry different works, this analysis is limited to the script and style of writing. Based on the palaeographical assessment and, in particular, on the remarkable similarity to the Leningrad Codex, as well as the historical discussion contained in these pages, one might be convinced that MS XXVIII(26) was produced in North Africa. However, the evidence remains scarce and one would need to compare MS XXVIII(26) to more manuscripts to reach some form of statistical significance.

5.5 Why is MS XXVIII(26) a good candidate for a digital edition?

MS XXVIII(26) is considered to be one of the treasures of the Biblioteca Capitolare of Verona. Its value as contemporary to Augustine as well as being the oldest surviving manuscript of his *De civitate Dei* make it an excellent candidate for careful conservation (as urged by Dr Alberto Campagnolo, the author of the conservation report of MS XXVIII(26) available in Appendix C) and for a digital edition for future generations to enjoy. Conservation is necessary in order to keep the materials constituting this precious object from degrading further; although the damage to some corroded leaves is likely irreparable, advanced digital imaging techniques, such as multi-spectral, could allow us to better read corroded text and may uncover text hidden under heavy mould (as previously seen in Figure 5.18). The digital imaging and restoration of MS XXVIII(26) would
be sufficient to foreground a documentary digital edition, one that would provide the first full transcription of this document ever. Such a digital edition, for which I have built a pilot demonstration with the best possible images at my disposal (see Chapter 6), would not only benefit the Biblioteca Capitolare in the way of publicity and conservation, but would also serve as a non-invasive and accessible research tool, allowing, for example, scholars to reuse the digital transcription in other (digital) critical editions or in stemmatological studies.65

The digital edition would in all likelihood be the only opportunity for researchers and the general public to view this manuscript. The restrictive opening hours of the Biblioteca Capitolare make it extremely difficult to study in the library (unsurprisingly, the very few people —rarely more than a handful— that can be normally seen in the main study room are either retired individuals or research students).66

5.6 Conclusion

The aim of this chapter is to bring together isolated studies concerning MS XXVIII(26), the oldest surviving manuscript of Augustine’s monumental De civitate Dei. Existing research tends to focus on specific aspects of its history: some, like Graham, have looked at its geographical and historical context; others, like Lowe, have contributed detailed palaeographical analyses. This chapter represents the first and only attempt to bring all this information together into one coherent whole.

The common thread connecting all of these studies is the debate on provenance. Most scholars interested in this codex have been torn between Italy and North Africa, two areas known to have been in close contact during the fourth and fifth centuries AD. Both learnt from each other and both influenced each other. How-

65However ambitious, given that the count of Augustinian manuscripts in European libraries (including but not only of De civitate Dei) is 15,000, see: http://www.augnet.org/?ipageid=2174 (Accessed: 11 June 2014).

ever, as Heine informs, ‘[Latin Christian literature] seems to have appeared first in North Africa’ (Heine, 2004, p. 131), a statement which also supports the belief that early Christian manuscripts such as MS XXVIII(26) were brought to Italy to escape the Vandal invasion (Lancel, 1999, p. 668). However, the lack of evidence questions existing scholarship. To this day, we do not possess definitive proof of provenance nor the exact production date of MS XXVIII(26). Such information might emerge with the digitisation and discovery of new manuscripts or perhaps even with a genetic test of the parchment.

But even if North Africa were definitively confirmed as birth place, the bigger question underlying all this research is: from which exemplar was MS XXVIII(26) copied? And could it be one of the first copies of De civitate Dei? It remains an open question and one that modern methods and resources might help us answer.

---

67 ‘Discovery’ is to be understood here not only as new findings in private collections or ancient caves (for example, the Dead Sea Scrolls - see http://www.britannica.com/EBchecked/topic/154274/Dead-Sea-Scrolls/233773/Discovery-and-description; Accessed: 30 September 2014) but also in existing library collections where cataloguing efforts may not be yet complete (See, for example: http://www.independent.co.uk/news/world/middle-east/nine-new-dead-sea-scrolls-discovered-in-israel-after-lying-unopened-in-storeroom-for-60-years-9189461.html; Accessed: 30 September 2014).

Chapter 6

A pilot, semi-diplomatic
documentary digital edition of MS
XXVIII(26)

This chapter describes ongoing work on the pilot digital edition of MS XXVIII(26) with a view to illustrating the rationale behind its conception and creation. This is the first known effort to reproduce MS XXVIII(26) digitally.

In 2011 this author purchased from the Biblioteca Capitolare of Verona a CD-ROM of digital images of MS XXVIII(26) that had been commissioned by the library in 2003. The CD-ROM cost 100 EUR and contains 258 JPEG images occupying 227MB of disk space (roughly 500KB per image). The embedded metadata of each photograph is the same and contains the following technical details:

- image dimensions: 2.857x2.207 (6.37Mpx); aspect ratio: 3:2 (1.50); bit depth: 8 bpp; color mode: YCbCr. Camera make and model: Nikon D100; lens: Nikon AF Zoom-Nikkor 24-120mm f/3.5-5.6D IF; aperture: F16; focal: 56.0mm (84.0mm); exposure: 1/13 s.

With the purchase of this CD-ROM, the spokesperson of the late rector of the library (Monsignor Alberto Piazzi) and I verbally agreed that these images could be published both online and in print so long as I provided the library with ten
copies of my doctoral thesis. In 2014, three years into the doctoral project and under the library’s new directorship of Monsignor Bruno Fasani, the agreement was annulled and with it the permission to publish the images online. The reason given for the sudden annulment was a conflict of interest that this project created with an upcoming regional initiative to re-digitise the holdings of the Biblioteca Capitolare of Verona.\footnote{See: https://www.regione.veneto.it/web/cultura/nuova-biblioteca-manoscritta (Accessed: 27 July 2018).} The amended permission dictated that I use an agreed number of images in the printed thesis only. Solutions I and my home institution proposed to mitigate the negative impact of the new agreement on my doctoral project were rejected by the new rector of the library. As a result, the documentary digital edition of MS XXVIII(26) described here cannot be published online with images but can be retrieved and viewed from the USB sticks submitted together with this thesis.

As is the case of the Catalogue of Digital Editions described in Chapter 3, in critically making a digital edition I aim to understand the complexity of digital resources in the fullest detail (Ratto, 2011) (Warwick, 2015). The expectation is that with first-hand, nuanced knowledge of the strengths and weaknesses of different technologies, of the time and skills needed to carry out specific tasks, of the dependencies tying different technological components to one another, of the trade-off between concept and implementation, of the balance between creation and use, and of the benefits of dissemination and publication obtained through the act of building a digital resource will, in future, help me formulate clear and realistic expectations of digital resources and of the people involved in their creation.

### 6.1 Methodology

In the planning stages of this digital edition, supported by the study described in Chapter 3, extensive research into technologies designed to facilitate manuscript transcription led to the discovery of the Digital Research Tools platform
6.1. Methodology

Based on the results of Chapters 3 and 4 with regard to transcription, I chose to produce a transcription in two formats, plain text (TXT) and XML-TEI (EpiDoc). Among the many tools and technologies listed in DiRT, three presented themselves as suitable and user-friendly candidates for the transcription of MS XXVIII(26): TILE, TextGrid and T-Pen. A full description of the advantages and drawbacks of each of these tools is beyond the scope of this chapter but a brief mention of their functionality is necessary to help to motivate the decisions taken for this digital edition. These three tools are free to use and support different modes and formats of transcription.

TILE, the Text-Image Linking Environment, was designed for image-based electronic editions and, as the name suggests, supported the creation of links between any token or word in a transcription and its exact location in the corresponding image. Unfortunately, however, a full working test for MS XXVIII(26) could never be completed as it was discovered that the tool had been developed by a short-term funded project, which was no longer maintained nor supported.

TextGrid provides a complete editing environment in which the user can gather and store files, edit and share texts, and even incorporate images. However, TextGrid’s multiple overlapping window-layout overcrowded my 13” MacBook Pro screen resulting in an unnecessarily cumbersome transcription process. This issue, coupled with performance delays of the software, called for an alternative, more stable solution.

T-Pen was tested by experimenting on one of the Latin manuscripts provided (St. Gallen, Stiftsbibliothek, Cod. 0002). The transcription process was facilitated

---

4For more information about TILE, visit: http://mith.umd.edu/tile/ (Accessed: 16 September 2014). The project was discontinued but the website is still running. The last blog-post dates back to July 21st 2011.
5A list of video tutorials and a video overview of TextGrid can be accessed here: https://www.youtube.com/playlist?list=PL_aqAE00s6w10Y9zkl5v7H9VvZByych5- (Accessed: 27 January 2015).
by the tool’s image zoning feature but was generally unreliable: indeed, aside from a number of browser issues, an unresolved bug in the system prevented the creation of new projects. For these reasons, T-Pen also had to be discarded.

The unsatisfactory experiments carried out with these tools cautioned me against the adoption of purpose-built tools and platforms of uncertain duration, favouring instead well-supported, user-friendly solutions that would also keep technological complexity to a minimum. As a result, I opted for an image viewer and a text editor to be used side-by-side: Mac’s native Preview application was used as an image viewer, while the transcription was carried out using the proprietary oXygen XML Editor. The advantages of oXygen over freely-available text editors such as Sublime Text or Text Wrangler include the support and validation of text encoded in accordance with the XML-TEI standard (see section 6.1.1).

Academic sustainability means keeping the content of the resource up to date. This may not be a problem with certain types of resource, for instance digital editions, which can retain their value for a very long period without any updating of their content. [...] But the problem of academic sustainability is necessarily less serious than that of technical sustainability, because resources will always retain at least some value whether or not their academic content is kept up to date, whereas if they are not updated technically they will sooner or later cease to be usable. (Robey, 2012, p. 149)

---

7 A list of known T-Pen issues is available on the tool’s Updates page, only visible to those who have a T-Pen account: http://t-pen.org/TPEN/ (Accessed: 20 January 2015).
8 The academic licence needed for the prolonged use of oXygen was provided by the Humboldt Chair of Digital Humanities at the University of Leipzig, my employer from May 2013 until December 2014. A 30-day free trial of oXygen can be downloaded from: http://www.oxygenxml.com/ (Accessed: 15 September 2014).
6.1. Methodology

6.1.1 Transcription

[... ] we expect digital editions to be searchable, malleable, collatable, quotable, analysable, which images alone do not allow. For these functions we need transcriptions, and if transcriptions are to be fully expressive, they must be encoded — that is what TEI is for. (Shillingburg, 2013, p.8)

Most of the visits to the library to access MS XXVIII(26) were made in 2011, at the start of the project, to assess the appropriateness of the images purchased as surrogates for the transcription work. Upon close inspection, the digital images of badly damaged folia did not prove suitable surrogates as zooming pixelated the image. My attempts to improve the clarity and readability (i.e., colour inverting and overlay) of the images with Adobe Photoshop was not successful. Supervised access to the manuscript itself was indeed granted but the short opening times of the library,10 coupled with my limited travel resources, made it impossible to generate a full transcription from the manuscript only.

The decision to use the Text Encoding Initiative (TEI) standard of XML mark-up was informed by the findings of the Catalogue of Digital Editions described in Chapter 3. Indeed, the large number of digital editions that encode text in XML-TEI is not only symptomatic of a well-established and supported technology,11 but TEI’s EpiDoc subset was found to be particularly suited for the reproduction of MS XXVIII(26). As the name suggests, while EpiDoc was originally conceived to encode inscriptions or epigraphs, its tagset lends itself well to other typologies of ancient text, including manuscripts.12

The transcription is evaluated against the Corpus Scriptorum Ecclesiasticorum Latinorum (CSEL) standard edition of De civitate Dei published in 1899, which

---

is freely available online in both PDF and XML formats. Other transcription aids used include the two online Latin dictionaries Whitaker’s Words and Perseus Latin Dictionary, as well as the print copy of Adriano Cappelli’s Lexicon Abbreviaturarum to help decipher abbreviations.

As mentioned in Chapter 5, MS XXVIII(26) is the oldest surviving manuscript of De civitate Dei; it is contemporary to St Augustine himself and preserves annotations of Pacificus of Verona. As such, it is deserving of a documentary digital reproduction to showcase all of its characteristics and peculiarities.

The transcription is semi-diplomatic insofar as it “represent[s] […] the visual features and signs of a particular document […] because they are significant” (Fischer, 2012, p. 78); it distinguishes between the traditional letter ‘s’ and the ‘long s’, however phonetically indistinct; it distinguishes between the letter ‘i’ and ‘j’, however phonetically indistinct; it breaks the scriptio continua by adding word-spacing; it keeps the original punctuation while disregarding spaces representing long pauses; it expands abbreviations; it points out omissions, misspelling, corrections and emendations; and, finally, it does not retain the distinction between large and small capitals.

Table 6.1 lists the Unicode characters used in the transcription of MS XXVIII(26) to reproduce abbreviations, special characters and scribal conventions.

---

13CSEL volume 40, parts 1 and 2
15Scriptio continua or scriptura continua is the practice of writing without word separation. For more information about scriptio continua, see the Blackwell Reference Online entry at: http://www.blackwellreference.com/public/tocnode?id=g9781444333275_chunk_g978144433327520_ssi-40 (Accessed: 29 January 2015).
16Large capitals indicate initials or letters that are larger in size that the rest of the text.
17As the official website of the Unicode Consortium states: “The Unicode Standard is a character coding system designed to support the worldwide interchange, processing, and display of the written texts of the diverse languages and technical disciplines of the modern world. In addition, it supports classical and historical texts of many written languages.” Available at: http://www.unicode.org/standard/standard.html (Accessed: 15 August 2015). The characters listed in the above table and used for the transcription of MS XXVIII(26) were taken from: http://unicode-table.com/en/ (Accessed: 15 August 2015).
### 6.1. Methodology

<table>
<thead>
<tr>
<th>Glyph(s)</th>
<th>Definition</th>
<th>Unicode</th>
<th>HTML Entity</th>
</tr>
</thead>
<tbody>
<tr>
<td>&amp;</td>
<td>Latin abbreviation for <em>et</em></td>
<td>0026</td>
<td>&amp;</td>
</tr>
<tr>
<td>Ψ</td>
<td>Latin small letter q with diagonal stroke on descender</td>
<td>A759</td>
<td>ꝙ</td>
</tr>
<tr>
<td>ρ</td>
<td>Latin small letter p with stroke through descender</td>
<td>A751</td>
<td>ꝑ</td>
</tr>
<tr>
<td>æ</td>
<td>Latin small letter ae</td>
<td>00E6</td>
<td>æ</td>
</tr>
<tr>
<td>œ</td>
<td>Latin small letter oe</td>
<td>0152</td>
<td>œ</td>
</tr>
<tr>
<td>Æ</td>
<td>Latin capital AE ligature</td>
<td>00C6</td>
<td>Æ</td>
</tr>
<tr>
<td>ṁ</td>
<td>Latin small letter p with flourish</td>
<td>A753</td>
<td>ꝓ</td>
</tr>
<tr>
<td>ṅ</td>
<td>Latin small letter long s</td>
<td>017F</td>
<td>ſ</td>
</tr>
<tr>
<td>ν</td>
<td>Latin small letter n with descender</td>
<td>A791</td>
<td>ꞑ</td>
</tr>
<tr>
<td>ꝕ</td>
<td>Insular g</td>
<td>1D79</td>
<td>ᵹ</td>
</tr>
<tr>
<td>‣</td>
<td>Combining ring above</td>
<td>030A</td>
<td>̊</td>
</tr>
<tr>
<td>☩</td>
<td>Cross of Jerusalem</td>
<td>2629</td>
<td>☩</td>
</tr>
<tr>
<td>☧</td>
<td>Chi Rho (monogram of Christ)</td>
<td>2627</td>
<td>☧</td>
</tr>
<tr>
<td>••</td>
<td>Three dot punctuation (archaic)</td>
<td>2056</td>
<td>⁖</td>
</tr>
<tr>
<td>†</td>
<td>Dagger</td>
<td>2020</td>
<td>†</td>
</tr>
<tr>
<td>Ṵ</td>
<td>Latin small letter Rum Rotunda</td>
<td>A75D</td>
<td>ꝝ</td>
</tr>
<tr>
<td> </td>
<td>Combining US above</td>
<td>1DD2</td>
<td>᷒</td>
</tr>
<tr>
<td>ḳ</td>
<td>Latin small letter Lezh</td>
<td>026E</td>
<td>ɮ</td>
</tr>
<tr>
<td>•</td>
<td>Middle dot (punctuation and abbreviation for -que and -bus)</td>
<td>00B7</td>
<td>·</td>
</tr>
<tr>
<td>−</td>
<td>Macron</td>
<td>0304</td>
<td>̄</td>
</tr>
<tr>
<td>••</td>
<td>Squared four dot punctuation</td>
<td>2E2C</td>
<td>⸬</td>
</tr>
<tr>
<td>N/A</td>
<td>Tridot composit (*, with bottom .)</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>

**Table 6.1:** Unicode characters used in the transcription of MS XXVIII(26) to reproduce special characters, abbreviations and scribal conventions.
As of 2017, Unicode does not support the following:  

- -NT ligature
- -UI ligature
- -UR ligature
- -US ligature
- -US ligature with long S
- -UT ligature

Where Unicode is not available, ligatures in this edition are encoded as `<hi rend="ligature">`. For example, the -NT ligature in word CREDANT in folio 10v, line 20, is encoded as:

![Figure 6.1: An example of ligature encoding](image)

Where a character is not yet available in Unicode, the XML `<app>` element is used to document the occurrence instead.

### 6.1.2 The critical apparatus

The critical apparatus generated by this digital edition follows the location-referenced method, whereby `<app>` elements appear within the text rather than in an allocated `<div>` at the bottom of the XML document. Each `<app>` contains a reading `<rdg>` element with a witness `@wit` attribute, which can contain one or more values. Different readings are encoded in separate `<rdg>` elements. Where manuscripts share the same reading, `@wit` can contain multiple values, as instantiated in Figure 6.2 below:

---

6.1. Methodology

Figure 6.2: The <app> provides the readings of manuscript BSB-Hss Clm 6267 (#Clm6267 in the XML), of Codex Sangallensis 178 (#Sang178) and of the standard CSEL edition (#CSEL40) in a single @wit attribute.

6.1.3 Line break

The line break element is used to represent lines in every folio. The attributes @n and @xml:id are used to specify the line number and create an anchoring point for text transformation purposes. Line breaks are also used to enable machine-readable references (Fischer, 2012, p. 81).

6.1.4 Abbreviations and corrections

As described in 5, the Latin script of MS XXVIII(26) occasionally presents horizontal markers above one or more letters. These markers are called macra (macron singular) and are used to abbreviate a word. Macra appearing at the end of a word usually replace a letter 'M' or 'N', while macra appearing in the middle of a work serve to elide a variable number of letters. It is only from context, meaning and palaeographical conventions that one can extrapolate the expanded form of the word. When encoding terms containing a macron, spacing is used to render the exact location of the abbreviation mark. For example, the macron at the end of the word ENIM (below) appears shifted to the right. To reproduce its precise location, a space must be inserted between the letter ‘I’ and the macron HTML character code, as follows:

Figure 6.3: The macron’s location in this example (word ENIM in folio 12v, line 8) is preserved by adding a space between the letter ‘I’ and the macron HTML character code.

Abbreviations were mainly adopted by scribes to save writing space.
If no space is added, the macron would appear above the letter ‘I’, thus misrepresenting the script. The encoding of macra in the Latin script is handled in XML via the <choice> element:

```
<choice><abbr>ENI &amp;#772;</abbr><expan>ENI&lt;ex&gt;M&lt;/ex&gt;&lt;/expan&gt;&lt;/choice>
```

Figure 6.4: Example of how the <choice> tag is used to encode the abbreviated word *ENIM*.

The <choice> element is also used to reproduce scribal corrections. Here is an example:

```
<choice><sic>ABS&lt;del rend="strikethrough"&gt;O&lt;/del&gt;RDA&lt;/sic&gt;&lt;corr&gt;ABS&lt;add place="above"&gt;U&lt;/add&gt;RDA&lt;/corr&gt;&lt;/choice>
```

Figure 6.5: The letter ‘O’ in the word *ABSORDA* has been deleted and replaced with a ‘U’ (folio 25r, line 13).

### 6.1.5 Errors: <sic>

I encoded overlooked or leftover errors and typos in a <sic> tag followed by an <app> element to provide the correct reading.

```
CAELUM ET <sic>TERAM</sic> <app><rdg wit="#CSEL40">TERRAM</rdg></app>
```

Figure 6.6: The word *TERRAM* below is missing one ‘R’ and is, therefore, tagged as an error (folio 14v, line 25).

### 6.1.6 Word break: <w>

Word breaks are encoded as <w> elements with the attribute @part. The values of @part used in this transcription are ‘I’ for *initial* and ‘F’ for *final*, as shown in Figure 6.7:
6.1. Methodology

6.1.7 Lost or damaged text: `<supplied>`

Text that was badly damaged, for example, is supplied inline via the `<supplied>` element. Damaged text in this manuscript is marked up as lost in the mandatory `@reason` attribute of `<supplied>`. Here is an example:

```
HIERUSALEM MATER NOSTRA AETERNA IN
```

Figure 6.8: Example of damaged text from folio 13r, line 16, and its corresponding encoding.

6.1.8 Addition and deletion: `<add>` and `<del>`

Scribal additions and deletions are marked within the text itself, not in the critical apparatus. Additions contain a `@place` attribute describing the location of the addition. For example, if the addition appears above a word or character, the value of `@place` will be `above`.

```
PRAEBENT UT <add place="above">PRO</add> EO QUOD
```

Figure 6.9: Scribal addition of PRO between UT and EO (folio 35r, line 9) appears above the main text and is thus encoded accordingly.

Deletions contain a `@rend` attribute describing the nature of the deletion. The strike-through value of `@rend` employed in the transcription of this manuscript is used to label crossed-out characters. Another type of deletion is expunction:
subscript and superscript (see Figure 6.10) dots were used by the ninth century scribe or corrector to mark characters to be deleted. The value of `@rend` used to represent expunction is `expunction`; `<del rend="expunction">` is occasionally followed by an `<app>` element describing the type of expunction, typically annotated below the letters to be deleted rather than above.

![Figure 6.10](image)

**Figure 6.10:** Example of expunction in folio 10v, line 27.

Finally, the value `erasure` appears where one or more letters or words were erased by the scribe. Where the erased letter(s) or word(s) have left a trace (Figure 6.12), I provide the original letters. If the letter(s) or word(s) have left no trace, these are made explicit through the CSEL 40 reading of the text in an `<app>` element.

![Figure 6.11](image)

**Figure 6.11:** Scribal deletion and encoding of the letter ‘A’ in *SEPTIMOQUAE* (folio 12v, line 21).

![Figure 6.12](image)

**Figure 6.12:** The pronoun *EIUS* is followed by what was once a second *EIUS*, which was subsequently turned into the verb *EST* by erasing the letters ‘I’ and ‘U’ and by adding a ‘T’ (folio 38r, line 6).
6.1.9 **Overlooked errors**

Orthographic errors in the text\(^{21}\) that have not been corrected by later scribes are encoded in a `<sic>` element followed by an `<app>` providing the reading of the reference text. The editor acknowledges the mistake and implicitly provides the correction via `<app>`. For example:

\[<lb\ xml:id="fo1_25r.26\" n="26\"/>TENEBRAS <sic>UULGATISSIMAS</sic> <app><rdg wit="#CSEL40">UULGATISSIMA</rdg></app>\]

**Figure 6.13:** The incorrect ‘S’ at the end of the word *UULGATISSIMA* has been overlooked. While the editor does not explicitly identify the mistake, the `<app>` element implicitly reveals the correct reading.

6.1.10 **Omission**

Omissions are implied in `<app>` elements documenting variant readings in CSEL. Below is an example of how the source text’s omission of *SENARIAM UEL* is acknowledged in `<app>`.

\[<lb\ xml:id="fo1_16r.28\" n="28\"/>>TUS</w> EST PROPTER SEPTENARIAM <app><rdg wit="#CSEL40">SENARIAM UEL SEPTENARIAM</rdg></app>\]

**Figure 6.14:** The encoded omission of *SENARIAM UEL*.

6.1.11 **Scholia**

As described in Chapter 5, MS XXVIII(26) contains many *scholia* or annotations in the margin. While the present digital edition does not record individual hands, the difference in handwriting clearly identifies the presence of multiple scribes and commentators. Scholia appear both in the margin and in the text itself (inline). They provide corrections, chapter markings or highlight particular sets of words. All scholia for this digital edition are encoded in a `<note`
type="scholion"> tag with the attribute @place to pinpoint their location on the manuscript page.

<note> elements are placed in or next to the line of text that is closest to their spatial location on the manuscript page. Computationally speaking, the location of <note> in the XML is irrelevant: an Extensible Stylesheet Language Transformation (XSL) can be written to extract <note> and place it wherever needed. From an encoding standpoint, however, placing <note> as close to its manuscript location as possible helped me browse the text more efficiently.

Unfortunately, not all scholia can be successfully deciphered. Placeholder comments <!--illegible--> (see Figure 6.15) are used in the XML to replace those annotations whose legibility is compromised by physical damage to the parchment and the low quality image resolution.

![Illegible annotation](image)

**Figure 6.15:** Illegible annotation in folio 13r, around line 4, and its respective encoding.

In such situations, illegible annotations can only be retrieved and read with the help of high resolution and/or multi-spectral imaging.

The pages of a codex were stitched together in gatherings of eight folia (sixteen if counting both sides).\(^{22}\) These booklets, known as quires, are documented in MS XXVIII(26) at the bottom of every eighth page:

6.1. Methodology

Figure 6.16: This leaf signature at the bottom of folio 38v stands for \textit{q(uaterni) III} (Roman numeral for 4, often also written as \textit{IV}), which tells the reader that this is the fourth quire in the codex. Signatures of quires one, two and three can be seen on folia 14v, 22v and 30v respectively.

6.1.12 Text encoding issues

6.1.12.1 Word break

As previously described, occasionally we come across abbreviated words split between two lines. For example, the word \textit{CONDERET} in folio 29r is split between lines 6 and 7 and contains a macron:

A validated encoding solution for this example is:

\[
<lb \text{ xml:id="fol_29r.6" n="6"} />...<choice><abbr><w part="I">CO \&#772;</w></abbr><lb \text{ xml:id="fol_29r.7" n="7"} /></choice><w part="F">DERET</w><abbr><ex>\text{N}</ex>DERET</abbr>\text{ex}</choice>...\]

Figure 6.17: An encoding solution for word breaks.

The same, however, cannot be applied to words split between folia, as illustrated in Figure 6.18 without breaking XML validation:
In these cases, an explanatory comment is used as a placeholder until an adequate solution is found.

6.1.12.2 Ligatures

As previously mentioned, the ligatures -NT, -UI, -US, -UR and -UT are not currently available in Unicode format.

At the time of writing, only the first two books of MS XXVIII(26), book 11 and 12, have been fully transcribed. The transcription of the remaining four books is ongoing.

6.2 XSL Transformation

The great advantage of electronic texts is that many decisions need not be made: the transcriber can include a wide range of information in the transcription but then choose how much of it to make available to readers or, better still, allow readers to choose for themselves how much of it they wish to see. (Driscoll, 2006, p. 258)

Driscoll’s words describe the essence of Extensible Stylesheet Language (XSL), a technology that transforms XML documents into HTML files for web presentation. XSL Transformations (XSLT) can be used to display an XML encoded text in different ways; for example, the editor or transcriber may want to create three separate transformations, each focussing on a specific aspect or layer of the encoded text: for instance, the first transformation might present the manuscript text stripped of editorial annotations, if any; a second transformation might
merge the text and the annotations into a single view; and a third transformation might be created to display the types of annotations the editor has made (e.g., philological or linguistic). These and many other presentation possibilities make XSL an ideal means of tailoring a digital edition to specific user groups and research interests. When combined with other computer languages, XSL transformations are also capable of generating dynamic and interactive presentations.\(^\text{23}\) By the same token, the customisable nature of XSL makes it a complex technology to master and one that requires the knowledge of another language, the XML Path Language (XPath), in order to query the XML text to be displayed in an HTML page.

For the digital edition of MS XXVIII(26) three sheets were produced to display the three different text parts of MS XXVIII(26): the introductory capitula, the main body of the text and the back matter. These three XSL sheets are very similar and could also be combined. However, as the layout of the text parts differs, for the sake of simplicity and to have better control over each component, I opted for three separate documents.

The stylesheets of MS XXVIII(26) display all XML (EpiDoc) mark-up, with the exception of the expanded form of abbreviations, which is yet to be implemented.\(^\text{24}\) Figure 6.19 shows an example of XSL code used to display the chapter number in the browser: the XSL code fetches the XML <note> tag containing the @subtype attribute with value ‘chapter’, and tells the XSL parser to display the contents of the tag as an HTML header (<h3>). The number of the chapter is retrieved via the <note>’s @n attribute.

\(^{23}\)JavaScript, for example - a popular web programming language used to animate static content. While the XSLT retrieves the parts of the XML text one wishes to be dynamic, it is JavaScript that manages the actual interaction functions. The two languages can be combined in an XSLT file.

\(^{24}\)The abbreviation ŚCÔ; for example, cannot be viewed in the browser in its expanded form SANCTO, which, as previously seen in section 6.1.4 is encoded in the XML.
Chapter 6: A pilot digital edition of MS XXVIII(26)

Figure 6.19: The first string of XSL code in the top of the screenshot is used to tell the XSL parser to retrieve information from the <note> element in the XML files; the second string of XSL code is used to tell the parser to retrieve the number of the chapter stored in the @n attribute of the <note> XML tag. The final HTML result, Chapter 1, is visible in the bottom part of the screenshot.

6.2.1 File management and storage

Consistent with the observations made thanks to the Catalogue of Digital Editions, I chose the versioning system GitHub to publish and store the ongoing transcription of MS XXVIII(26). Furthermore, the transcriptions are licensed under a Creative Commons Attribution ShareAlike 4.0 International License (CC BY-SA).25

6.3 The digital edition website

The 'Infrastructure' column of the Catalogue of Digital Editions documents a wide range of technologies, platforms and software being used to build digital editions and their hosting websites. This large diversity attests to the complexity of digital editions, and to the differing needs and resources of their creator(s).

In order keep technological complexity and obsolescence to a minimum, the hosting website of MS XXVIII(26) being developed rests on HTML pages generated by XSL transformations of the XML manuscript transcriptions, and on the Bootstrap framework to support responsive and mobile-friendly brows-

---

6.3. The digital edition website

Bootstrap integrates HTML and CSS, two of the core technologies for building webpages, with JavaScript, a dynamic programming language typically used for web interfaces. By employing these core technologies and avoiding databases, content management systems or web-publishing platforms (e.g., Omeka), I strive to minimise web maintenance efforts.

Figure 6.20 illustrates the file structure of the digital edition of MS XXVIII(26). The constituting file-sets of the capitula, the back matter and of individual books are stored separately for better file management and navigation.

![Diagram of file structure]

**Figure 6.20:** The file system underlying the digital edition of MS XXVIII(26) stores all files making up the capitula, back matter and individual books in separate folders for better .

Figure 6.21 shows the `index.html` file, that is, the home page of the digital edition of MS XXVIII(26):

---

29 Web-publishing platforms, as useful as they may be, require the installation of various components and are subject to regular updates. Omeka, for example, can only be downloaded if its hosting machine meets certain requirements, and the download package includes Apache, PHP5 and MySQL—all components regularly maintained and updated by their developers. Omeka also supports the installation of plugins, whose functionality might break if not compatible with the latest release of PHP or MySQL. For more information about Omeka, visit: http://omeka.org/
A top navigation menu lists the eight textual divisions in the manuscript; a click on the first menu item, *Capitula*, opens a drop-down menu with a list of all folia constituting that division, as shown in Figure 6.22.

Each folio page is divided into two regions, text and image. The text region on the left-hand side contains two tabs, one for the manuscript text and one for the editorial notes/critical apparatus (see Figure 6.23). The image region on the right provides a dynamic rectangular zooming frame to enlarge the text (see Figure 6.24).

A bottom 'Previous/Next' pagination allows the user to click through each folio in sequential order.
6.3. The digital edition website

Figure 6.22: A top navigation menu lists the eight textual divisions in the manuscript; clicking on any one opens a drop-down list of all folia making up a particular division. Alternatively, users can use the bottom ‘Previous/Next’ pagination for a sequential reading of the text.

6.3.1 Bootstrap

As previously mentioned, Bootstrap is an HTML, CSS and JavaScript framework developed to create responsive and mobile-compatible websites. As the number of smart hand-held devices grows and as the findings of Chapter 3 reveal, it is desirable that digital editions be mobile-friendly for wider access. As a framework, Bootstrap comes with a default CSS style and a default set of JavaScript scripts. The <head> of the <html> document illustrated in Figure 6.25 contains links to Bootstrap’s CSS stylesheet, bootstrap.min.css, and to two JavaScript files, jquery.min.js and bootstrap.min.js. These files regulate the aesthetics,

---

30 The Ofcom Adults’ media use and attitudes 2015 report reveals that 2/3 of UK adults own a smartphone or tablet and that six in ten (61%) adults now use a smartphone to go online. These statistics are located on pages 28-29 of the report, available at: http://stakeholders.ofcom.org.uk/binaries/research/media-literacy/media-lit-10years/2015.Adults_media_use_and_attitudes_report.pdf (Accessed: 4 January 2018).
Figure 6.23: The two-tab layout of the textual region of the edition page provides both the transcribed text and automatically generated notes listing variants encountered in other witnesses.

animations and responsiveness of Bootstrap HTML pages.

Figure 6.25: Default CSS and JavaScript Bootstrap settings must be added to the `<head>` of the website’s HTML documents. Customisations are possible and must be added as separate files in the `<head>`.

If satisfied with the default settings, the developer or editor need only interact with HTML in order to create the layout of the website. The digital edition of MS XXVIII(26) uses Bootstrap’s default settings; the only customisation is text colour, as illustrated in Figure 6.26 (green and superscript for additions; magenta and square brackets for supplied text).
6.4. Users, use, usefulness and usability of the digital edition of MS XXVIII(26)

Owing to the restrictive image rights previously outlined in Chapter 5, a user study of the pilot edition of MS XXVIII(26) proved challenging, if ineffective, to arrange. Had the study been possible, the search for suitable candidates for user-testing would have begun at relevant Departments and Centres at University College London, including the Information Studies Department,31 the Depart-

ment of Greek and Latin, the Centre for Digital Humanities, the Centre for Editing Lives and Letters and the Centre for Medieval and Renaissance Studies. In light of the mixed audience anticipated for this digital edition (students, scholars and general public alike), the differing scope and expertise of these departments would provide the feedback diversity required to tailor the resource to as wide a range of needs as possible while informing its future development. Additionally, should a change in permissions allow the online publication of the digital edition, I will propose it for review in the Review Journal for Digital Editions and Resources (RIDE) to increase its visibility and draw the interest from a wider academic community.

The digital edition of MS XXVIII(26) is addressing both active and passive users mainly through its simplicity, and is designed to accommodate the three uses described by Dillen and Neyt (2016, p. 3) both from desktop computers and mobile devices thanks to Bootstrap technology. The landing page of the digital edition provides all the information necessary for it to be included in the Catalogue of Digital Editions as described in Chapter 3. Each page or folio can be printed in colour or black and white thanks to the dual mode of marking editorial intervention. The transcribed text is being made available via GitHub in both TEI(EpiDoc)-XML and TXT formats to enable faceted reading (i.e., the ability to read different layers of detail, e.g., expanded abbreviations) and computational analysis. The types of search anticipated include the study of particular features (e.g., abbreviations), the scholia, and of quotations and other textual parallels to other witnesses or works. At present, these searches can only be carried out on the source-code (both XML and TXT) but, as mentioned in Chapter 5, the implementation of search functionality within the digital edition itself is planned as a future development.

6.5 Intellectual property

While the Biblioteca Capitolare holds the rights over the digital images of MS XXVIII(26), the text of MS XXVIII(26) itself is in the public domain, as deliberated by COMMUNIA\textsuperscript{36} and by the Copyright Statute of Italy (Art. 27),\textsuperscript{37} the country in which the manuscript is kept.\textsuperscript{38}

The transcriptions that make up this digital edition are my original work.\textsuperscript{39} They are published under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. This means that the work is open for reuse as long as it is adequately credited and shared under the same conditions. The selection of this particular licence was motivated by my commitment to open access and open source research, and the additional NonCommercial clause was dictated by the Biblioteca Capitolare’s explicit request that no commercial gain be made from this research.\textsuperscript{40}

Were the digital edition of MS XXVIII(26) to be finally published online, it would be released under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

\textsuperscript{36}Available at: http://www.publicdomainmanifesto.org/italian (Accessed: 4 January 2016).
\textsuperscript{39}See, for example, Art. 85 quater 1 of the Italian Copyright Statute. Furthermore, Dr Maria Xalchou, author of a 2014 edition of Codex Vindobonensis phil. gr. 65, writes: “Whoever carries out the transcription and scientific study of a manuscript holds the Intellectual Property of the work.” See: https://www.academia.edu/7796987/The_copyright_and_the_Intellectual_Property_and_Right_of_the_edition_of_the_Codex_Vindobonensis_phil._gr._65_which_is_a_Hellenic_Manuscript_found_at_the_Austrian_National_Library_in_Wien (Accessed: 4 January 2016).
\textsuperscript{40}For a general overview of the relationship between digital editions and copyright, see Dillen and Neyt (2016).
6.6 Future work

One thing is clear: the more we put into a text, the more we can get out of it. (Driscoll, 2006, p. 261)

The future developments envisaged for this digital edition include the addition of indices of people and places mentioned in the text, the addition of search functionality, the creation of a Zenodo Digital Object Identifier (DOI) to facilitate citation, a thorough assessment of the digital edition’s conformity to web usability standards (e.g., colour-blindness), to FAIR Data Principles, and to the findings from the user survey.

6.6.1 TRAViz: visualising text variation

One of the benefits afforded by the digital realm is visualisation. Digitally, a static stemma codicum can, for instance, be transformed into a dynamic tree or graph for close inspection and manipulation of textual anomalies or of analogies and differences between witnesses.43

[...] I do not consider it a requirement for an electronic edition to display textual variation in an apparatus or in any other way, for in my view such display has a project-specific purpose. But [...] the study of textual variation —where it appears— is an essential part of the research involved in creating an electronic edition. Even a reading edition cannot be made without a serious examination of the textual variants. (Vanhoutte, 2006, p. 163)

An ongoing experimental development for the digital edition of MS XXVIII(26) is testing the variant visualisation capabilities of TRAViz (Text Re-use Alignment Visualization) to bring together different readings of De civitate Dei.44 TRAViz is

---

41 Available at: https://zenodo.org/ (Accessed: 2 May 2018).
43 For more information about the typologies of text variant graphs, see Schmidt and Colomb (2009).
openly available under a Fair Academic Licence (FAL) and repurposes some of the functionality of the analogous CollateX and Stemmaweb tools.\footnote{CollateX is software used to read and compare two or more versions of a text. For more information, visit: http://collatex.net/ (Accessed: 15 September 2014). Stemmaweb is a collection of tools for collated text analysis. For more information, visit: http://stemmaweb.net/ (Accessed: 15 September 2014). Another manuscript comparison tool is Manuscript Comparator, available at: http://prototypes.openscriptures.org/manuscript-comparator/ (Accessed: 26 August 2015).} Its ease of use and design make TRAViz an effective means of comparing variant readings. In fact, users need only add a plain text transcription of the manuscripts they wish to visualise to the index.html file provided in order to view the collation, as shown in Figure 6.27.

Figure 6.27: To visualise variant readings of a text, the TRAViz user need only add plain text transcriptions of the manuscript witnesses to the HTML file provided; in this example, I have added five different texts to compare.

For demonstration purposes, the test graph visualisation for the digital edition of MS XXVIII(26) brings together the following readings:

- MS XXVIII(26) [early 5th century], Verona (Italy);
- Cod. Sang. 178 [mid 9th century], St Gallen (Switzerland);
- BSB-Hss Clm 6267 [9th century], Munich (Germany);
- Cod. Lat. 121, late 15th, Budapest (Hungary).\footnote{Available at: http://www.e-codices.unifr.ch/en/list/one/csg/0178}

\footnote{Available at: http://daten.digitale-sammlungen.de/~db/0003/bsb00039815/images/index.html?id=00039815&fip=ewqrrseayaxdsydxsydxsydxsydxsydxsydxsydxsydxsdydxsydxsydxsdydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsydxsyo\

\footnote{Available at: http://www.corvina.oszk.hu/corvinas-html/hub1codlat121.htm}
When opened in a browser, the `index.html` TRAViz file displays the variant graph visualisation shown in Figure 6.28.

Figure 6.28: Excerpt of book 11, chapter 3 of *De civitate Dei* (folio 9r, lines 8-9 of MS XXVIII(26)). For every word in a graph a pop-window lists all witnesses in which the word appears.

In this visualisation, each reading or witness is assigned a colour. Readings can be isolated by clicking on the corresponding coloured thread. As Figure 6.28 shows, for every word in a graph a pop-window lists all witnesses in which the word appears. The size of a word in the visualisation is indicative of its occurrence across the witnesses under scrutiny. For instance, the words *ANIMO*, *MENTE*, *SENTIUNTUR* and *IPSE* are larger in size because they appear in all five readings; conversely, *HAC* is much smaller as it only appears in Cod. Lat. 121.

By the very nature of the plain text format, should I need TRAViz to visualise both the first text of MS XXVIII(26) and that as amended by the 9th century hand(s), I would have to provide two separate plain text transcriptions. In reality, therefore, while TRAViz gains in user-friendliness, the amount of data pre-processing required to prepare plain text files may prove unfavourable. Indeed, to visualise an XML-encoded text with TRAViz, all transcription files must be stripped of their XML mark-up, which, depending on the depth of the encoding, may become an onerous task.

---

Furthermore, TRAViz also raises concerns with regard to scalability, in that, to the best of my knowledge, there is no evidence of it being used to chart variants in large manuscript traditions, such as that of *De civitate Dei*. It remains to be seen whether, for example, twenty different (complete) witnesses of *De civitate Dei* (amounting to circa 5,500,000 running words),\(^{50}\) which would have to be digitised (if not already) and transcribed, would have an impact on the loading time of the JavaScript code that controls the visualisation. Another, perhaps more troubling, concern is the unreliable support offered by the developer(s) of TRAViz. My numerous requests for support went unanswered even though the tool is still being developed, used and presented at conferences.

Since TRAViz, new collation solutions have emerged: the *Edition Visualization Technology* is gaining momentum thanks to its automatic handling of XML transformation and variant collation,\(^{51}\) while the tabular approach of *PyCoViz* is particularly useful for the visualisation of large manuscript traditions.\(^{52}\)

### 6.7 Conclusion

This chapter has served to describe the creation and ongoing development of a pilot digital edition of MS XXVIII(26), and to consider its use, usefulness and usability with respect to the recommendations put forward in Chapter 4. The restrictive reproduction terms imposed on this work by the Biblioteca Capitolare in Verona do not allow me to publish this digital edition online. Nevertheless, this proof-of-concept serves as a reminder of the effects of poor digitisation on scholarship. Indeed, the work described in these pages would have significantly benefited from an improved set of photographs, which I will continue to pursue together with permission from the Biblioteca Capitolare to publish the digital edition on the web.

---


Chapter 7

General conclusions

This thesis describes the creation of a pilot digital edition of MS XXVIII(26), the oldest surviving manuscript of Saint Augustine’s (354-430 AD) monumental *De civitate Dei* (The City of God). Also known as Manuscript *V[eronensis]*, MS XXVIII(26) dates back to the early fifth century AD and is housed in the chapter library of Verona, Italy. As contemporary to Saint Augustine himself, it is a particularly treasured object of study. This thesis reassesses extant research about this manuscript, collecting information about its disputed provenance, historical context, materiality, tradition, and conservation. In doing so, it investigates how the manuscript can be best reproduced as a digital edition by way of two surveys designed to better understand how digital editions are respectively being created and used. In so doing, this research provides a better understanding of how the digital editions being built by the Digital Humanities community are meeting the needs of their users.

Chapter 2 outlined the history and theory of (digital) editing with a view to capturing the evolution of this practice over a period of two millennia. Chapter 3 followed with the description of my ongoing effort to catalogue extant digital editions, which today number in the hundreds, to identify the *status quo* of the production of digital editions. In Chapter 4, 242 of these catalogued digital editions were compared against the results of a web survey in which 218 members
of the Digital Humanities community expressed their needs and expectations with respect to these projects, thus clearly delineating converging and diverging points between production and consumption. Chapter 5 reassessed all existing knowledge of MS XXVIII(26), the manuscript chosen as a case study for the creation of a pilot digital edition, in line with the findings gathered from Chapters 3 and 4. The ongoing development of this pilot digital edition is described in Chapter 6, which also reflects upon its use, usefulness and usability.

### 7.1 Research contribution

The research value brought by this thesis is threefold. Firstly, it contributes a growing list of digital editions in the form of a large body of evidence, which can be queried, analysed and visualised to concretely grasp the development of digital scholarly editing as a field. Starting out as a personal study tool, this data-set organically grew into a collaboration with the Austrian Centre for Digital Humanities at the Austrian Academy of Sciences and is now acknowledged by numerous international initiatives in the Digital Humanities. The conclusion deriving from this research is that there is no community consensus over how digital scholarly editions should be built but that standards do guide creators through the implementation of their project. Digital editions appear to be moving away from a traditional print publication layout and toward more experimental data-centric and reuse formats.

Secondly, this thesis provides the largest survey yet to have explored the consumption of digital editions. With an unprecedented number of participants from the Digital Humanities community, the ‘Expectations of Digital (Textual) Editions’ survey compares 218 responses against 242 digital editions present in the Catalogue of Digital Editions, resulting in a list of recommendations for funders, libraries and creators of digital editions alike toward the development of more user-centric projects. Specifically, to promote quality, sustainability and usefulness, both creators and funders of digital editions must devote human and
7.2 Consideration of method

A large part of this research involved practical work in order to concretely and critically assess the theory of digital scholarly editing. This process of “Critical Making” employed an intentionally mixed research method, quantitative and qualitative, to fully understand the dichotomy between theory and practice, and between creation and use of digital editions (Ratto, 2011).

As detailed in Chapter 3, the data collection for the Catalogue of Digital Editions qualitatively assesses projects through a process of descriptive cataloguing, and quantitatively measures practices of creation by aggregating as many exemplars as possible. This content analysis raises issues of reliability and validity, in that while taking an objective approach, the task of cataloguing does not rely on an inter-rater agreement but on the efforts of a single individual. Reliability concerns may, in turn, question the validity of the data collected, which, again, while seeking the highest accuracy is only valid if the digital edition is not subsequently altered. As a result, the Catalogue of Digital Editions is regularly revised to account for any changes and additions that may compromise the validity of the data. In the absence of an automated means of revision, the Catalogue currently relies on the manual efforts of its team and users.

The qualitative research into MS XXVIII(26) made almost exclusive use of printed literature. Considerable time, effort and resources were needed to combine manuscript viewings with desk research at the Warburg Institute in London,
for which access had to be specifically granted. Indeed, very little information about MS XXVIII(26) is digitally available and the effect this paucity had on the research timeline had not been anticipated.

It is striking that humanities scholars seem to take the availability of automated catalogues and searchable back runs of periodicals for granted […]. (Prescott, 2012, p. 19)

The transcription of MS XXVIII(26) is also a qualitative task and one that was initially greatly underestimated. I had planned to submit this doctoral thesis with a complete transcription (254 folia) of MS XXVIII(26). However, as the manuscript transcription progressed, the plan had to be readjusted according to the prolonged timeline and the resources available, thus downsizing the project to a pilot digital edition showcasing the capitula and one of the six books constituting MS XXVIII(26) (corresponding to ca. 90 folia of manuscript text).

### 7.3 Dissemination of results

Portions of this thesis have already been published in peer-review journals and volumes or are currently in press. Parts of Chapter 5 can be read in an article entitled TRAViz: A Visualization for Variant Graphs, published in 2015 in the journal Digital Scholarship in the Humanities (formerly LLC). Chapter 2 was published in 2016 in Digital Scholarly Editing: Theories and Practices, an Open Book Publishers volume edited by Elena Pierazzo and Matthew Driscoll. Finally, Chapter 3 was accepted for publication in late 2018/early 2019 in the Special Issue on Evaluation of Digital Cultural Resources of the ACM Journal on Computing and Cultural Heritage (JOCCH). A journal publication is also planned for Chapter 5.

### 7.4 Looking back

The research presented in this thesis has been a hard but exciting journey of discovery and growth, both professional and personal. The learning curve was steep but the critical making of both the Catalogue of Digital Editions and of the
digital edition of MS XXVIII(26) was essential for a comprehensive and deep understanding of the field of digital editing in general and of the complexity and challenges of digital editions in particular. The process of discovery of digital editions through the creation of the two related resources mentioned above has been an extremely positive and enriching experience. Looking back, the only, damaging, mistake was failing to secure written permission from the Biblioteca Capitolare in Verona to use the digital images I purchased in the pilot digital edition.

7.5 Recommendations

The following recommendations are based on my own experience and are intended for those who are planning their first digital edition or analogous digital cultural resources.

7.5.1 Time

Creating a digital edition of a text is a time-consuming and complex effort.

The pilot digital edition described here began with a literature review of digital (scholarly) editing to better understand the state of the art in the field and to learn more about the key components of a digital edition. The time spent to read through four-decades’ worth of history of digital editing and to reach the level of notional confidence needed to proceed with the creation of the digital edition of MS XXVIII(26) was approximately one year. The information gathered from this initial step helped to recognise the benefits that the digital edition might bring to its intended audience and to the visibility of the manuscript itself.

The literature review informed the transcription of MS XXVIII(26). While the initial intention was to produce a complete transcription of the 254 folia constituting MS XXVIII(26), once the transcription was underway it become clear that only a fraction of the manuscript could be digitally reproduced within the doctoral time-frame. Indeed, the transcription of the first book of MS XXVIII(26)
(book 11) took approximately one year (as a part-time activity, corresponding to roughly six-months of full-time commitment); such a span might initially appear excessive but it includes the time needed to scope out and test the appropriate XML tag-set, repeated revisions of the tag-set to meet newly encountered features as the transcription progressed, as well as the meticulous comparison of the transcription against a standard edition of the text to enforce high academic quality. A complete transcription of MS XXVIII(26) is expected to take three years; the transcription of variant readings would prolong this task.

The preparation of a website or web space to host the digital edition is also a lengthy process. When creating a website, one must sift through a plethora of available technologies and publishing platforms to find the solution that best fulfills the needs, scope and sustainability plans of the digital edition. The more complex the website, the greater the effort needed to ensure that technological upgrades do not compromise the inner-workings and cross-compatibility of the various components of the website. The website of the digital edition of MS XXVIII(26) was set-up and perfected over a period of two months.

As Chapter 3 has revealed, users of digital editions expect to find comprehensive documentation to help them understand the editorial and technological choices, as well as the research scope of the project. Writing up adequate documentation can take weeks.

### 7.5.2 Images

The survey described in Chapter 3 reveals that a primary user expectation of digital editions is the inclusion of digital images of the source materials or texts under study. It follows that one of the first tasks in the creation of a digital edition is to secure written permission to publish existing images or to arrange for new ones to be taken. Moreover, if at all possible, images should be published under open licences to allow for download and reuse in other projects.
7.5.3 (Infra)structure

Another result of Chapter 3 points to users’ general preference for data over aesthetics or functionality. In other words, a digital edition need not be overly elaborate or cutting-edge but should rather focus on providing clear and easy access to the raw data. The more accessible the data, in terms of both licensing and formats, the more likely it will be repurposed and/or enhanced by users. Furthermore, the simpler the digital edition, the easier its maintenance over time.

7.5.4 Training

Many scholars in the field of digital editing have pointed out that owing to their complexity, digit editions are seldom the work of a single individual. The research described in this thesis corroborates these assertions, inasmuch as it demonstrates that digital editions are best achieved through collaborative work, harnessing the skills of multiple individuals. Flaws in the digital edition can occasionally be traced back to the absence of the particular skill required to fix them. Whether working alone or as part of a team, researchers who build digital editions should be prepared to follow training courses or sessions to help them understand and operate the various components.

7.5.5 Users

The research underlying Chapter 3 shows that user studies in the field of digital editing are few and far between. Lacking an understanding of what users seek from a digital edition can lead to its neglect. If possible, creators of digital editions should explore the needs of their target audience, as this information can help shape the project and potentially save valuable resources.

7.5.6 Dissemination

Dissemination and regular activity have been shown to reduce the neglect of a resource (Warwick et al., 2008, p. 389). Indeed, while the digital edition of MS
XXVIII(26) could be published online, the regular addition of fresh content to the Catalogue of Digital Editions described in Chapter 2 (advertised on Twitter and ResearchGate) is helping to build up a following and to increase visibility, which in turn results in citation and in the use of the resource. Regular dissemination and content creation also contribute to improving search-engine indexing and ranking, pushing a resource to the top of relevant research results. Creators of digital editions are encouraged to upload content in small, regular instalments not only to help garner continued user interest but also to increase the chances of the resource being discovered, popularised, cited and used.

This thesis set out to identify the *status quo* in the field of digital (scholarly) editing from both a creation and user standpoint to deepen my understanding of the complexity of digital cultural resources and of digital editions in particular.

The investigation began with a literature review to learn more about the history and the evolution of the theories and methods of digital (scholarly) editing. With that foundational knowledge, I began to catalogue a large number of digital editions to empirically assess and tease out the state-of-the-art. The tabular format used to collect and store the data was used to statistically infer patterns and trends, which were then compared against a web survey involving more than 200 users of digital editions from the Digital Humanities domain to better understand to which extent creation practices are meeting the expectations of those who consume these resources. The preliminary results obtained from these combined studies serve as the first large body of tangible evidence of the field’s failure hitherto to address users from the Digital Humanities domain; the diverging points identified between creation and use of digital editions can be used by practitioners to help shape future projects. Indeed, these results are informing my own pilot digital edition of MS XXVIII(26), which remains to be user-tested but seeks to become a useful and usable reproduction of the oldest surviving manuscript of Saint Augustine’s *De civitate Dei*.

As well as outlining creation and use in digital (scholarly) editing, I have also
translated my first-hand experience in building a digital edition into a number of recommendations for those embarking on analogous projects and for (funding) institutions who support their creation.

Diversity, complexity and technology are pushing digital editions into new and experimental directions, inviting them to take new shapes and to constantly evolve to keep up with the times. For creators of editions of text, the challenges have never been greater.
Bibliography


Bulatovic, N., Gandt, T., Romanello, M., Stiller, J., and Thoden, K. (2016). Usability in Digital Humanities - Evaluating User Interfaces, Infrastructural Compon-
BIBLIOGRAPHY


Ciula, A. and Eide, Ø. (2014). Reflections on cultural heritage and digital human-


BIBLIOGRAPHY


Greetham, D. (2013). A history of textual scholarship. In Fraistat, N. and


Hsieh, H.-F. and Shannon, S. E. (2005). Three Approaches to Qualitative Content
BIBLIOGRAPHY


BIBLIOGRAPHY


4037288?uid=3738296&uid=2129&uid=70&uid=4&sid=21181955345741 (Accessed: 
31 March 2013).


London.

Montague, E. and Xu, J. (2012). Understanding active and passive users: The ef-
eaffects of an active user using normal, hard and unreliable technologies on user 


Nielsen, J. (2006). The 90-9-1 Rule for Participation Inequality in Social Media 

Ore, E. (2009). “...They Hid their Books Underground”. In Deegan, M. and Suth-
Ashgate.


ator work in automated process control - a job design case study in a control


Appendix A

‘Expectations of Digital (Textual) Editions’ Web Survey

A.1 Survey email announcement

Expectations of Digital (Textual) Editions: A Short Questionnaire (20 questions, 15 minutes max).

Do you use and/or build digital (textual) editions? If so, please consider filling-in this short questionnaire, which aims at collecting information about what users expect or want from a digital edition.

QUESTIONNAIRE URL: https://opinio.ucl.ac.uk/s?s=48797

The questions build upon the feature list provided by the Catalogue of Digital Editions [https://github.com/gfranzini/digEds_cat and https://dig-ed-cat.eos.arz.oeaw.ac.at/], and the answers obtained from this questionnaire will be examined against the editions currently contained in the Catalogue. The information you provide will help us compare the user needs of the community with the digital editions that have been built by the community.

The questionnaire contains 20 questions and is completely anonymous. We don’t ask for demographic information such as age, gender, ethnicity
or religion.

The compiled results of the survey will be made available online via the Catalogue of Digital Editions websites and the questionnaire’s institutional address (University College London). They will also be discussed in Greta Franzini’s PhD thesis.

The questionnaire should take no more than 15 minutes to complete and closes on 30th April 2017. Please share it with colleagues and friends who might be able to contribute!

For further information about this questionnaire or about how the data will be used, please contact Greta Franzini at g.franzini.11@ucl.ac.uk

Thank you very much for taking time to fill-in this questionnaire. We truly value the information you provide.

Greta Franzini, Prof. Melissa Terras, Simon Mahony

A.2 Answers given to Question 3
Question 3
How would you describe your subject area (e.g. I'm a Classicist but my current focus is on Classical Philology rather than Classical Archaeology or History):

Medieval historian
Byzantine studies; epigraphy
Medievalist
Digital Literary Studies with a focus on French and Spanish literature.

Digital Humanities with some more profound knowledge in History and Philosophy
Lexicologist (medieval french), philologist, developer of dictionary writing systems
I'm in Law, but my PhD was a diversion into Information Studies.

History
Classicist - Latin literature - focus on texts, textual analysis. Also interested in ancient philosophy, particularly Stoicism, and what happens to it in the Roman period. Plus I do classical reception.

Literary studies & history of ideas
Former senior TV professional, BA in education, now on a masters course MSc Digital Education.

Digital media studies
Classical, but with a passion for modern English and German Literature
Late Medieval and Early Modern Literature
I'm a medievalist researcher.

Classics, Latin literature

Corpus Linguistics, Discourse Analysis and Translation Studies

I'm an Arabist, but currently work in a library

Digital Classicist
Classics, Ancient Greek and Latin, Archaeology, History

History of ideas, in particular philosophy and language sciences, Eastern and Central Europe
I am specialist in Classical and Medieval Latin, and work on History of Scientific (better: Medical) Texts

I'm a Computing Scientist with a background in Ancient Greek.

Computational Linguistics for Literature

French literary history

Computational linguistics applied to humanities data
Classically trained in philology; research focused on late antique, medieval, and Byzantine texts in Greek, Syriac, and Latin.

Medieval Historian
I'm a Computer Programming with interests in classical studies
I'm a PhD student in English philology, with a focus on manuscript studies and English historical linguistics.

Medieval history of law

Epigraphy

I'm an Americanist with interested in race and representation in contemporary literature. I am also interested and actively researching digital and critical pedagogies.

Digital Humanities

ENGLISH LITERATURE (POETRY AND NOVELS)

Literature person with a focus on fantasy, science fiction and children's literature.
Early modernist (primarily Renaissance England) with a focus on book history and digital humanities

Medieval history of the Near East

One of my main foci is the modelling and localization of all models related to the reconstruction of copy histories or textual evolution given medial transitions (oral->chirography->print->digital)
Medieval language and literature

Romance philology

New Testament

Medieval studies, Germanic languages, Old English language and literature.

As a PhD student, my subject area is architectural anthropology, architectural history and historiography and material culture. As an HE manager, my subject area is Digital Humanities and strategy.

Greek and roman history

I'm a literary modernist.

I'm a Classicist but my current focus is on Classical Philology

Early Modernist with interest in DH

Textual Scholarship with a focus on modern German literature

Italian Studies and Digital Humanities

I'm a classicist. I used to work on Augustan Latin literature, but now my focus is on DH and digital scholarly editions.

I previously studied Classics (in particular Classical Philology), while now I'm focusing on Digital Humanities

DH

Historian (Early Modern Times), Digital Humanist, Library/InformationScientist

Medieval English literature and philology

Religious Studies, textual scholarship on Syriac

I am a textual scholar and a digital humanist with a specialty in English literature.

Classical Archaeology and History

Classical Philology

I'm a Classicist with a focus on Classical Philology (Late Greek Epic).

I'm a Byzantine historian, though I was trained as a Classicist.

classist-generalist-with an interest in epigraphy

I'm interested in Medieval Latin literature, comparative literature and digital philology (especially of medieval texts).

Corpus linguistics and historical sociolinguistics

Mediaeval Studies

I'm a Classicist, not interested in philology but researching Civilizations, History, Archaeology and creation and applications of 3D models

Digital philology

Agriculture

My background is In Romance Philology and genetic philology. I researched on digital philology but currently I'm interested in how technology affects learning and teaching.

Manuscript Studies

Romance Philology

Codicologist and palaeographer specialising in Anglo-Saxon Manuscript culture now working in DH/archives/information studies.

I studied library and information science and Romance languages and am currently managing Digital Humanities Projects.

Medieval Medicine

I am researching the intersection of art and science.

Biblical Studies (New Testament, Textual Criticism, Greek)

My focus is on library and information science, digital libraries, digital humanities, digital cultural heritage

Late medieval-early modern England--approx. 1400-1550.

greek law, greek epigraphy (ancient)

I'm a Classicist, primarily a philologist, with interests in Greek religion and the reception of Greek literature by early Christians.

I'm a digital humanist in an English department who focuses on digital literature and adaptations.

Languages and literature

I am a citizen scientist involved in a variety of digital humanity and natural history crowdsourcing projects and am also interested in crowdsourcing as an an academic discipline although I am not affiliated with any academic institution.

art history

19th and 20th century American literature, book history, history of photography
I have a background in Social Sciences but I am currently focusing on History
Sociology

Shakespeare Studies

Digital Humanities/Digital Archiving

Medievalist, History of exototic animals in the Middle Ages

Classicist with a focus on literary studies

Generally work in English and Scottish Language and Literature but also work across broader Humanities subjects

Classicist specialised Roman Law and Ancient History

Shakespeare studies, digital philology

Midern Literatures

My subject area is the liturgical medieval chant

Medieval History/Art History
digital philology

I'm a digital librarian with a background in English literature.
I'm a philosopher with responsibilities in digital editing.

British Literature

Area Studies, Early Modern Printing, East-Asia

Classicist with a focus on Classical Philology.

History of daily life

Linguistics

eighteenth century English literature and culture

Digital History

I am an engineer, my main background is on automation. Worked mainly in the power system and the cyber security sector. Recently got an MS degree on classical letters, with a thesis on a digital edition of Archilocus.

I'm a Classicist with a focus on Ancient Greek Literature.

I am working on folklore and Digital Archiving

I'm linguistically oriented classicist using all kinds of sources, text-object-visual

I tend many topics connected in both disciplines. Computer Sciences is a privileged part of our culture and Philosophy is originated in and compromized with the culture.

Ancient History, Military History

Scholar of modern and contemporary Italian Literature, which I want to study by applying Digital Humanities

Surgery, History of Medicine, History in general

I am a historian focusing on all aspects of the sixteenth century

historical metrology

Digital Humanities, creating digital editions

medieval legal historian

I'm a Sinologist, my main work consists of creating and maintaining digital texts and related tools

I'm a historian of the ancient Mediterranean with philological and epigraphical training

Text in the digital age

I'm working in the field of Patristics

Textual critic

I'm a medievalist (late medieval-early modern)

I'm a Classicist but my current focus is on Classical Philology

Theatre director specializing in Early Modern English drama. One-time professor of theatrical literature. Lecturer and speaker focusing on Early Modern English drama with a focus on Shakespeare and his contemporaries.

Classicist focused on language pedagogy.

Digital Humanities, combining Phd in English with a career in IT.

Medievalist by training (manuscripts and historiography), with substantial work experience in humanities computing / digital humanities (textual criticism, codicology, American literature). Now an independent scholar in medieval MSS again while newly employed in the software industry. I can't tick "non-academic professional" after so many years with universities. :)

Particular interests in linguistics
Classicist working in medieval philosophy
Composition and Rhetoric, and Visual Rhetoric
History
I am an historian of the transmission of texts and sources in the Middle Ages
architectural historian and musicologist developing database-driven web-portals; focus is on long-time persistence of digital
data (mainly, but not restricted to) in the humanities
Renaissance Italian literature and culture
Mostly library and information studies, which tends to branch out into a lot of different subject areas depending on the
research
I am focussed on Military History, primarily regarding both World Wars, and ego documents (letters and correspondences,
diaries, autobiographies).
Historian (Diplomats, paleography), now encharged in a digital humanities position
Ancient Near Eastern Philology / Hittitology
English Language and Applied Linguistics
Long c18th History
Renaissance studies, mainly renaissance intellectual history/philosophy
Literary historian currently working on theatre history
Medieval manuscripts and book history, interested in physicality over text.
I'm an early modern historian focusing on late sixteenth/early seventeenth century cheap print.
Early modern history and literature
Early modern english scholar
I teach British literature before 1800 as well as a variety of writing classes. My research focuses on accessibility and
usability in digital environments.
English Literature; Children's Literature; Drama
Food History
English Renaissance Drama
medieval history, focus on editing medieval latin texts
I'm a historian by training but my research is specifically related to digital scholarly editions - so I generally say my subject
area is digital humanities
I'm an IS professional working in cultural heritage digitization and textual scholarship
Hispanic philologist but working on digital humanities
Mediaevalist, Jewish studies
I'm a philologist, though my current focus includes education and modern languages education.
Early American literature
English Literature
Digital Humanities
I am trained in Editorship of Polish Literature and my dissertation was in 16th century translation of the Book of Psalms, but
I am more interested in meta research about editorial studies rather than in particular literary epoque. Lately I've been
studying User Research.
Corpus and Computational Linguistics in Greek
I'm a historian using and editing medieval sources to analyse medieval society
Philosophy
Medieval literatures and cultures
I'm a medievalist who focuses on textual criticism and translation in late medieval England.
Latin médiéval, Histoire médiévale
Literature scholar with focus on contemporary experimental literature
Classicist
I work on all facets of early modern letter-writing, from linguistic to material.
Journalism, history, new media
Medieval English Literature
Early Medieval language literature and culture
Biblical studies
Early American Literature
research infrastructure (especially for DH)
Classicist doing archaeology and history.
I'm an archaeologist but currently working on legacy data in museum collections
Professional Writing, with a bend toward the technical: IA, procedures, instructional text
Medievalist who has a deep background (previous career) in programming and databases
I am a digital archaeologist focusing on 3D visualisation in prehistory
data-driven sociologist of culture and science
Textual Scholar, Digital Humanities & Caribbean Literature/Culture
Classics, specifically Roman history and religion
Computer scientist focused on interaction design with a background in DH related to textual criticism (Donne, Cervantes, New Testament and a few others).
Digital art history
Classicist: Late Antiquity and Byzantine poetry. Philology. Text crit
I'm a Classicist working currently on Classical Philology
Musicology
Classical Philologist: mostly textual criticism, mostly Latin, mostly verse.
I'm an italian philologist and I'm focused on documentary texts written in twentieth century, specifically on private letters composed during the first world war.
Social history of the late Ottoman Middle East, media history, conceptual history
medieval history, auxiliary sciences (diplomastics, palaeography, codicology, sigillography)
Archivist & historian working as a professional archivist in a university research centre.
library information science + DH + classics
Classicist working on a mixture of literature, art, myth and "factual" material.
Historian, with focus on interdisciplinary field known as "history of the book"
Specialist in material culture of the ancient Near East with a focus on demonology
History
I'm a philologist, primarily a Celticist, but I branch out into Old English and Latin too.
I take care of technology and data modelling of digital textual editions and also maintain data and web frontends
Art history, D(A)H
Classical Philology
Bureaucracy
A.3 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.
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?

- **university de Caen**: double encoding of critical apparatus: XML-TEI and notes dynamic interface scientific liability provided by the critical edition + images see also

- **e-codices**: Excellent high quality images and good metadata. I can do the rest.

- **Welscher Gast Digital**, [http://digi.ub.uni-heidelberg.de/wgd/](http://digi.ub.uni-heidelberg.de/wgd/) Among many things: - fine grain information about the graphic system (allographs, abbreviations, punctuations), and a configurable interface; - downloadable sources in XML/TEI; - synoptic visualisation of the manuscript transcriptions; - CC BY SA licence; - DOI for citability; - digital facsimile.

- **The Electronic Beowulf** [http://ebeowulf.uky.edu/](http://ebeowulf.uky.edu/) - searchable and easy to cite

- **DC3 edition** pioneered by Hugh Cayless, [https://www.youtube.com/watch?v=n-pWR3iAyE](https://www.youtube.com/watch?v=n-pWR3iAyE)

- **The Digital Fragmenta Historiorum Graecorum** meets my needs. I would like an edition made with its principles and including the critical apparatus, maybe done through the Perseids platform to avoid copyright infringements.

- **LOEB Classical Library**. The edition is very basic, but it has the text, the translation, the notes and search options. Brilliant if you do not have immediate access to the physical books.

- **The digitizing program of the Vatican Library and of the Bayerische Staatsbibliothek**

- **The Digital Fragmenta Historiorum Graecorum** meets my needs. I would like an edition made with its principles and including the critical apparatus, maybe done through the Perseids platform to avoid copyright infringements.

- **Epigraphic Database Roma Epigraphic Database Heidelberg Epigraphic Database Bari EAGLE**

- **The Electronic Beowulf** [http://ebeowulf.uky.edu/](http://ebeowulf.uky.edu/) - searchable and easy to cite

- **Welscher Gast Digital**, [http://digi.ub.uni-heidelberg.de/wgd/](http://digi.ub.uni-heidelberg.de/wgd/) Among many things: - fine grain information about the graphic system (allographs, abbreviations, punctuations), and a configurable interface; - downloadable sources in XML/TEI; - synoptic visualisation of the manuscript transcriptions; - CC BY SA licence; - DOI for citability; - digital facsimile.

- **Epicodic Database**

- **The Electronic Beowulf** [http://ebeowulf.uky.edu/](http://ebeowulf.uky.edu/) - searchable and easy to cite

- **Welscher Gast Digital**, [http://digi.ub.uni-heidelberg.de/wgd/](http://digi.ub.uni-heidelberg.de/wgd/) Among many things: - fine grain information about the graphic system (allographs, abbreviations, punctuations), and a configurable interface; - downloadable sources in XML/TEI; - synoptic visualisation of the manuscript transcriptions; - CC BY SA licence; - DOI for citability; - digital facsimile.

- **Epicodic Database**

- **I would cite my own digital edition here, so I will abstain from answering this question]**

- **Text, apparatus, textual analysis, witness (with images), commentary, bibliography, translation, data search, download**
There is none

I like lots of editions but at the moment particularly http://digi.ub.uni-heidelberg.de/wgd/ it has weaknesses (i.e. too much information everywhere, I feel confused) but I think the way they capture information in facsimiles and make them in integral part of the edition is somehow innovative (i.e. http://www.wgd.materiale-textkulturen.de/illustrationen/motiv.php?m=1).

http://www.beckettarchive.org/

When I was studying Classics I used the Perseus Collections of Greek and Roman Materials. I found it useful because it includes a wide range of texts, and, as well as the original Latin and Greek texts, it has many translations, so sometimes I could find a text that wasn’t available in the library or wherever I was. Also, I could search for a word within a work to find parallels (if I wasn’t in a library with access to the online Thesaurus Linguae Graecae).

The Shelley-Godwin Archive, although not for materials within my period, looks like the type of edition I would use, providing images, different transcription views, and downloadable XML. It also hybridizes the single-text edition with the archive, allowing the study of multiple objects within the same platform. Within my own period of the English Middle Ages, the new online version of the Electronic Beowulf would be the gold star.

Not satisfied with currently available options.

The Online Variorum of the Origin of Species. It shows a synoptic view of variation which helps to understand the timing of the changes. It is a pity that a mobile App has not been developed.

Perseus Digital Library.

Hm. The only example I can think of isn’t public, since it’s one I’m building myself (and thus is uniquely designed to cater to my needs). In my experience, Byzantinists aren’t always enthusiastic about this kind of technological approach, so I’m the only one I know who’s working on something of this nature. My main concerns in making this edition for myself were 1.) to document where and how this particular scribe diverges from the canonical version of a given scriptural commentary, and 2.) to clarify what exactly this manuscript contains, to better understand how/why the copyist put it together in this way. It’s also quite handy to have something I can do corpus analysis with!

Perseus comes closest for teaching and some research, I am probably not fully aware of what is available.

Beowulf online by Kevin Kiernan Dnate's Comedia by Peter Robinson Skaldic Project

The way that Bible Software (Accordance, Logos) has implemented digital original languages is quite useful. Has mobile access, lexicons, apparatuses, notes, etc. Very accessible and the platform is already there.

I can’t

Oxford’s Holinshed Project (http://www.cems.ox.ac.uk/holinshed/): diplomatic transcription to plain (not TEI) text; parallel texts of different editions; annotations; easy, reliable copy-and-paste.

IApH 2007: It contains the texts and all the necessary metadata

Most volumes contained in the Biodiversity Heritage Library provide good functionality for my needs. It would be helpful if more of the individual articles in some of the journals contained in the Biodiversity Heritage Library had DOI’s attached to them. The would enable easier citations. But the fact I am able to download images and pages for reuse makes this platform the most useful for me in my work.

http://www.walkerart.org/collections/publications citation, images, context, primary resources, multimedia

Edgar Allan Poe Society of Baltimore is more like a complete Poe resource; it’s very technologically but very thorough and extremely reliable as far as texts go. I use it for my own research and for teaching.

Perseus Odyssey

Olive Schreiner Letters Online www.oliveschreiner.org

Internet Shakespeare Editions from the University of Victoria. It maintains very high standards of accuracy and gives its stuff away.

www.cn-telma.fr; www.unicaen.fr/scripta1; http://www.gasconrolls.org/fr/

The Digital Vercelli site (http://vbd.humnet.uni.pi/i/beta2/#) and the tool that powers it are very good. It’s possible to view facsimiles and transcriptions side by side, or switch to a purely facsimile or transcription view. It has decent search facilities, a good facsimile zoom tool and just feels like a very well put together system. (Note that I have no involvement with this site or the tool).

Bibliotheca Iuris Antiqui (BiA)

Folger digital texts: reliable modern editions, in a variety of formats (e.g. txt, tei xml). These can be downloaded or searched on the online platform.

I consider the Modern Journals Project a very good digital edition. Technically, it's not an edition, but an archive--but this is why I like it so much. Each text provided is clustered natively within a meaningful context (a particular medium during a particular artistic movement). This allows specialists to run comparative analyses or historical/material readings on these digital editions of particular magazines while allowing students or the general public simply to access the texts. What this means for me is that the digital edition can play two roles: providing open-access availability to the public and providing researchers with new avenues for research.

https://tei2016app.acdh.oeaw.ac.at/ provides an api to retrieve data, is open source (code of the application) and open data

http://www.catulusonline.org/CatulusOnline/index.php meets my needs because it is openly accessible and of a very high quality in terms of scholarship and availability of materials.

In addition to the facilities provided by the Perseus Digital Library and Perseids, platforms like EVT and tools like the Classical Text Editor seem to provide most of the needed features, on a case by case basis. My MS thesis (available on request) provides a peek-up of the digital edition I have designed, in view of providing a support to teaching Ancient Greek literature to university students. This integrates repositories, tools and other facilities available from the above.

Not now

Free Courses On line (Coursera, Open University, Stanford, MIT, etc.)

Petarchive (Petrarch's Reum Vulgarium Fragmenta) because its great functionality (images; text; writing)

Just about anything published by Tanner-Ritchie. Totally free of scanning problems; material difficult to find in libraries in my part of the world. Aside from the photographed hands and illegible gutters often encountered, my greatest problem is 19th c works scanned at such a low resolution that the footnotes (in 6 pt type) are entirely illegible. The worst material is the Early English Books collection.

https://www.briefedition.alfred-escher.ch – clear editorial principals, good user interface

Something like the Early English Laws project is useful because it lets you put things side by side and also copy things in a plain text format to use in other forms/mark-up outside of the browser.

My work at www.kanripo.org and www.mandoku.org tries to build editions that fit the bill.

PHI Greek Inscriptions, though only for personal research since it lacks translations. It provides easy and quick access to primary source evidence that I can later look up in published but less accessible versions when necessary, but online allows me to formulate conclusions and engage in scholarly conversations in a timely fashion.

Dickinson Classical Commentaries (http://dcc.dickinson.edu) - a little more interactivity/customization would be appreciated, but it is currently good enough to use in place of a book for those students who are comfortable using an online text. Actually no, I haven't found any edition yet that is available both as ePub and website with an API, so that it allows me both to just read and enjoy it on a tablet (or even as print) or use it as teaching material, and that allows me to create interesting new entry points by reusing elements of it through its API (or download facilities). I'm sure there must exist some, but apparently not in my areas of interest.


The online Froisart

Icelandic Saga Map http://sagamap.hi.is/is/ - Complete corpus - Well defined corpus - Reliable corpus and curation - No feature overkill. It does what it does well. - Utilises IT to augment the texts with features that would not be possible in print and with enhance our ability to appreciate and understand the texts. In this case, finding places across all texts, linking places to a map and linking the map to locations in the corpus.

I'm not sure I have found one that meets my needs for teaching. I teach research writing and digital editions of textbooks could help students see the inquiry process from idea to publication. It's all too formulaic and simplistic at this point to be more helpful to me than the materials I have gathered through past student examples.

Corpus Corporum (Philipp Roelli) - because the corpus is extremely rich, but it does not have the functionalities I would expect for a digital CRITICAL edition (the quality of each works depends of the edition used to be put in the corpus).

Verhandlungen des Deutschen Reichstags und seiner Vorläufer (Session Reports of the German Reichstag and its Precursors, http://www.reichstagsprotokolle.de/index.html) - offering reliable and scholarly fully re-useable information as full text resource - allowing cross-search through various resources

http://gams.uni-graz.at/context:srbas?mode=projekt Jahrrechnungen der Stadt Basel You can interact with the material by collecting items and calculate prizes

Syltistics by Paul Simpson. Easy to get to and navigate. Does what it says on the tin.

No.

Norton Shakespeare online provides (some) performance information alongside fully-edited text, glosses, and other information; allows comparison of different editions/formats (so interactive and flexible, with good user interface)

Depends on how you define edition I guess. Christoph Flueler has talked about digitized mss as digital editions and I am swayed by his arguments. It is probably not what you are considering here but. So going along with CF I would say that the best editions for my use include e-codices, OPenn, and the Digital Walters.

I currently use EEBO, which has mostly everything that I need, but transcription of text is occasionally unhelpful/unavailable. The ability to search via STC nos is invaluable, as is the ability to study the text closely. The variant spelling function is useful, as is the fuzzy function.
I love the Folger Digital Texts Shakespeare (and associated API) http://www.folgerdigitaltexts.org/, http://www.folgerdigitaltexts.org/api

TEAMS medieval texts: explains what the source is, gives basic commentary, is clear it's not the ne plus ultra scholarly edition but is free, online, handy for quick reference when libraries are closed or the library I belong to doesn't have the scholarly edition available.

I don't know of any that meet my teaching needs--computers in classrooms still not universal. But folger digital editions are pretty good.

Nothing immediately comes to mind.

Cambridge Ben Jonson: it has proper scholarly apparatus and accurate texts.

I use Open Source Shakespeare a lot. Though it has problems, the search functions, line numbers, and provenance make it a reliable and trusted online edition of the complete works. I also use the Shakespeare app for iOS, which functions more or less the same as OSS, except that it has more text (includes the poems and a couple more plays). I use these two sources most frequently when I'm away from my physical editions at home and need to check a reference, or to copy-paste a large passage for analysis in my work. Though I'll always check those digital editions against my Ardens or my Norton, being able to access them on the fly is super handy. I most often use them for the search function when I'm trying to find a particular passage or for language analysis (how often does X word get used, etc.). When preparing texts for performance, I will generally start with the OSS text and reformat/cut from there, but PlayShakespeare.com and the Folger digital editions are also good sources for that. MIT Shakespeare is terrible and should be stricken from the internet.


I cannot recall at the moment.

Briefedition Leopold Wilhelm

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

The Becerro Gallicano of San Millán: easy access to the source material and a range of search tools

http://www.oxfordscholarlyeditions.com/ an example of a simple, clever, and efficient interface

When earlier in the survey the idea of user interactivity and a "complete" experience was introduced, I realized how rare that is. I like the text part of the Bayeux Tapestry edition http://www.bayeuxtapestry.org.uk/index.htm and the user experience of the Letter of Prester John (though it is missing some vital elements of an actual scholarly edition) http://scalar.usc.edu/works/prester-john/index. Something that combines these would be ideal.

van Gogh Letters, http://vangoghletters.org/vg/ - good survey, quick orientation, smart navigation/hierarchisation of information - transparency (working principles, definitions etc.)

Not really: every edition is a compromise. Some are excellent at accessing the texts but don't provide images (e.g. http://www.livesandletters.ac.uk/bodley/bodley.html ); others may have more contextual information and indexing and images, but require more clicking to actually get to the texts and/or images (e.g. https://www.bessofhardwick.org ). I think there's probably a trade-off between adding bells and whistles and making something readily accessible.

Beowulf Keenan and to a much lesser extent Bayeux Tapestry Foyis ( because this did not use actual images ) both good intros to scholarship around the two texts for U/G s with no / little experience of studying MSS

Not really: it's hard to come by into an edition that provides every source material needed to replicate the result.

There are very few digital editions in my field. One I have used recently is Memorable Days: the Emilie Davis Diaries. Valuable because makes available a text previously hard to access, including to the manuscript. It has drawbacks too.

Abby Covert's _How to Make Sense of Any Mess_ (http://www/howtomakesenseofanymess.com) turned out to be an interesting case for me this past year: I've worked with the physical paperback, the ebook, and the website versions of this test, and it turned out that the ebook was the _least_ useful of the three because of it's degraded random access.

I can't identify one yet, except perhaps my own work to be published in the next year. I struggle to find digital editions which really address social history concerns.

https://elotroalex.github.io/ed/ (Full disclosure: I built it)

The Perseus Project combines open access texts with an accurate parsing/dictionary feature.

Oxyrhyncus at Oxford

Dante Alighieri: Commedia. A Digital Edition ...if it was open access Some strong points: High-resolution colour digital images of manuscripts and easy zooming Collation of variants linked to a map Go directly to any view of any line in any witness from a menu

Working on my own because nothing available does what I want them to do.

The online Froissart project https://www.hrionline.ac.uk/onlinefroissart/

hmmm, difficult choice, those of perseus/oppure and digiliblt are good approximations

Can't think of a specific document, but essentials include: - PAGE NUMBERS!!!! (so many editions leave these out, which means I can't offer them for student use as I expect students to cite page references) - An accessible footnote/endnote system - Searchability - Ability to change size of fonts

Yellow 90s Online (The Yellow Book)

No

nothing

Only partly. The Confessio edition is well done but interaction with the text is limited and it isn't downloadable. Granted it oriented towards the general public.
A.4 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.
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?

Text input

Clear instructions about the LOD possibilities. How easy is it to identify/link to parts of a text? How deep can I go? To an individual word? What is the desirable minimum - perhaps MSS numbers, locations, names?

no

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. (One of the questions above asks about browsing/searching and API in one question.)

imho, a lot of DEs focus too much on the “D” (too many functionalities because they are simply possible, but without reflection if they are of use) and not enough on the content of the edition: the aspect of being “scholarly” needs to be stressed, if not we lose about a century of scientific tradition and best practices. Users need information on the ms. (date, language, etc.), on the content of the text (via glossary or linking to dictionary entries), etc. All that is sign of a good scholarly edition on printed form needs to be sign of a good digital scholarly edition.

Simply a text on a page and the ability to CTRL+F is often the best way for me to find things - although I would like to have a more efficient search function to find e.g. all the uses of a particular word in a particular text!

Producers of digital editions should refrain from developing new formats or software applications, but try to use existing and wide-spread tools to increase usability.

Avoid online-only versions whenever possible, or at least provide an offline-readable version like a PDF.

Linking to other projects using the same base texts

I think it is important that digital editions are done using standardised markup use as much OpenSource Software as possible, to make everything reusable. Editions done with proprietary file formats that only work in specific programs will help nobody in the long term.

How can we have a digital edition tailored to various needs in one place?

Some editions are not user-friendly at all. Many e-books in my library are just scanned versions of the paper publications - no search options, no links, no possibility to highlight text, etc. In some cases you also need to download a specific app for reading those books, which is really uncomfortable.

I need access to the individual ingredients that make up a complex edition. A series of witnesses bound to a complex, collated edition is difficult to use to extract the readings for only MS A, for example. However you decide to do your edition, make available as individual files the simple components that helped make it.

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

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.

There is always room for improvement, but digital editions are important for me so that I can find information that my library will not have and so that I can bring information into the classes I teach. If the directors or editors of digital projects make it easier to work with, all the better!

I’m an aberration because I now do 99% of my work on my iPad or Galaxy Tab, including in class tool demonstrating. I rarely use a PC, unless I need 30” screen, or proper command prompt work. So I use, and recommend to students, only tools which work across mobile platforms.

I would be interested in how central the question is to born-handwritten chirographic age texts, if we can represent a text the witnesses of which show variation in one curated maybe emended version in order to spare the readership the time and effort to dive into that jungle. Is it a central question or are digital editions rather for experts. A digital edition of a born digital text is also something completely different, which I think requires other possibilities and interactive text creation/adaptation modules such as found in crowd sourced poetry.

Maybe the question of citability and sustainable identifiers might be of interest too.

You haven’t mentioned two important issues: ease of use of the user interface, and the citability of the edition itself.

Cross edition compatibility and interoperability – beyond the tokenism of proselytising for the TEI.

- 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.

No.

Embarrassingly simplistic introduction to digital resources in classics. I think I need to attend some programs and I’d love a private PhD consultant in classics!

The knowledge of textual criticism and language tools should be mandatory.

I’m all for editions doing everything by the book and providing reusable data, metadata and whatnot. But honestly as a scholar I don’t care at all. What I really need is an interface that lets me search the text and check the results in an ergonomic way. Sometimes, very rarely, I might need the XML source or whatever, but 99% of the time I’m happy just reading.

If the edition makes phrase and key word searches easy, it will be usable. If it simply highlights every instance of every word in the search, it isn’t very helpful.
See above. As a Manuscripts specialist, good quality images are essential. I want images I can access freely, download, and then do something with (e.g. DigiPal/VisColl).

What about preservation?

Again, for my purposes, I value the actual historical/literary content over what I understand might be data value as words in corpus.

permanence, discoverability

Digital editions are not just of interest to and production by those in literature and cognate areas!

A doubt about the definition of the digital edition: a thesaurus and a bibliography for access to digital sources are considered digital edition?

complete access to the TEI code (not only the generated html on a website) and the editorial guidelines are very important to me.

I would also find it useful if there is an option on the digital platform of the digital edition where one can have a notebook and space for creating collections of images.

I disagree with Sahle’s view that facsimile editions are not scholarly.

I believe that most of the features listed in previous questions are very important! So this means for me that I can appreciate projects that roll out a simpler first version and then systematically roll out richer features.

clear and open license, provide a manual so that users can actually RTFM

Knowing that the edition reflects the judgment and experience of one or more editors adhering to a specific methodology is important to me.

You may contact me on: alberto.stefanini@virgilio.it

Sustainability over time is an important issue. In my view this is best guaranteed by “also” providing a downloadable offline version of the edition which is still functional.

I think it is very important to keep track of updates and different versions of editions. One option could be the “archived copies” that, for instance, the Bayerische Staatsbibliothek has of same web-sites.

There are many other things that could make a good digital edition for other kinds of texts and corpora, links to supporting and related material, timeline, full text search, ability to integrated in other systems with api's, ability to access in multiple formats such as read on web, download as epub/pdf etc.

The ability to assign portions of the text for reading, analysis, etc. Sometimes repositories of resources are not indexed and I have to tell students, “Search Using Charts to Visualize Data.”

Important scholarly editions I use go back to the 16th century, e.g. Daniele Barbaro’s annotated editions of Vitruvius. I can use these still after 460 years. How long will users be able to use digital editions of today? This is the crucial question for me, because I cannot accept an answer like “20, maybe 50 years”. That would not be worth the effort, time, money, person years usually spent for any edition. We should find a general solution, immediately!

The ability to export parts of the document, as well as the ability to annotate in ways similar to traditional marginalia just as easily (ideally in a form that can be exported or synced) is critical.

Having a special output format or design for PDF would always be great, but is a very rare feature. Retrieved information could be restructured and given in a format that directly names the resource(s). Would be perfect if one could include such extracts into a systematically material collection.

Changes / Updates should be traceable.

Nope.

A powerful search facility is always useful (for single words and variants)

They are hugely beneficial; one size fits all not always going to be best though!

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.

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

When I think about a digital edition, I in general feel that its UI it underserved. I used some of the editions that are available online during classess with students and I could see that the more dated the website looks, the less likely they are to pay attention to what I’m saying. The fact that the material is prepared well should be a given—but without the presentation, we’re lost. Digital editions should take various groups of interest into account, not only the scholars, but also students and perhaps even casual readers.

sustainability, discoverability, quality control - the very foundation for any publication on the web
We need to move towards more open access to both key scholarship in the fields of Digital editions as well as access to the MSS themselves.

There needs to be a repository for single-work editions — not only large scale, multi-text editions like the Walt Whitman Archive which tend to replicate canonical hierarchies. "Just Teach One" on Common-place is in my opinion the model for the future of digital editions: small scale, free, not dependent on grant funding, decentralized, a passion project — but making immediate impact on the field.

Generally editions have focused on philological concerns/apparatus. Very little conversation has been given to non-literary uses, in particular legal and financial documents. This opens up a completely different field of what should be tagged as meta data.

Bibliographies.

Classics digital editions often have poor browsing features within the texts (the Loeb Classical Library, for example). The ability to easily scroll through texts and browse within them is crucial.

An ODD file documenting the TEI use would be precious.

I want a 'create-your-own-text' edition where users can select entries in the on-line app.crit. to promote to the text, demote other choices, add their own conjectures, and save their own mini-editions for printing or later retrieval and possible re-editing. Quot Lectores, Tot Editores!

There is a need for explicit licensing for the every component of an edition as well as for machine-actionable (bibliographic) metadata and stable URLs.

good idea this survey!

As a researcher outside of academia who works in the arts and humanities, I need digital editions of texts that function more like traditional print publications with perhaps a few additional features, such as embedded videos or hyperlinks. I have not yet seen any digital publications in my field, art history, that are especially successful at reconciling the need for high quality images with a successfully designed publication integrating text and images. I tend to just use PDFs and then find high-quality images elsewhere.

The problem is which audiences to target. To attract linguists you need full markup - but then the text is perhaps better placed in a corpus in the first place. I have a text I want to make available online myself to a) give greater/more varied access to the MSS than is possible in a printed book and b) to make a high quality, interactive text available to the public and scholars simultaneously. I am still trying to decide the best approach from both an impact and time/resource-economical point of view.

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

A standard for critical appratuses which show more than the reading variants is desperately needed.

User-friendliness

4 / 4
A.5  **Full survey report: complete and incomplete responses.**
### Question 1
You are a:
Question 2
Your background/interest is in:

<table>
<thead>
<tr>
<th>Choices</th>
<th>Absolute frequency</th>
<th>Cum. absolute frequency</th>
<th>Relative frequency by choice</th>
<th>Relative frequency</th>
<th>Cum. relative frequency</th>
<th>Adjusted relative frequency</th>
<th>Cum. adjusted relative frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Humanities (Archaeology, Anthropology, Classics, History, Languages and Literature, etc.)</td>
<td>270</td>
<td>270</td>
<td>83.08%</td>
<td>77.59%</td>
<td>77.59%</td>
<td>90.91%</td>
<td>90.91%</td>
</tr>
<tr>
<td>Applied Sciences (Computer Science, Mathematics, Physics, Engineering, etc.)</td>
<td>24</td>
<td>294</td>
<td>7.38%</td>
<td>6.9%</td>
<td>84.48%</td>
<td>8.08%</td>
<td>98.99%</td>
</tr>
<tr>
<td>Interdisciplinary (please specify the disciplines)</td>
<td>31</td>
<td>325</td>
<td>9.54%</td>
<td>8.91%</td>
<td>93.39%</td>
<td>10.44%</td>
<td>109.43%</td>
</tr>
<tr>
<td>Sum:</td>
<td>325</td>
<td></td>
<td>100%</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Not answered:</td>
<td>51</td>
<td></td>
<td>14.66%</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Average:</td>
<td>1.26</td>
<td>Minimum:</td>
<td>1</td>
<td>Variance:</td>
<td>0.39</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median:</td>
<td>1</td>
<td>Maximum:</td>
<td>3</td>
<td>Std. deviation:</td>
<td>0.62</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Total answered: 297

Last choice text input
- Digital Humanities
- Digital Humanities
- Digital Education
- Political Science, Sociology, Digital Humanities
- DH
- Linguistics, Biology, Computer Science
- DH; World Art; Material Culture
textual scholarship
- Digital Humanities
- Digital humanities, digital education
- DH
- Book History and DH
- Library and Information Science, Digital Humanities
- Natural history, digital humanities, crowdsourcing, open access, copyright and reuse
- Social science esp sociology
- Medical Humanities
- Philosophy and Computer Sciences
- Philosophy & Computer Sciences
- Medicine, History
- Medicine, History
Medicine, History

Computer science, textual scholarship, science and technology studies

Humanities, IT

Information Studies

Digital Humanities

User Experience Research, so bordering on Social Sciences

Library and Information Science

social sciences, media studies
Question 3
How would you describe your subject area (e.g. I'm a Classicist but my current focus is on Classical Philology rather than Classical Archaeology or History):

- Editor of Latin charters.
- Medieval historian
- Byzantine studies; epigraphy
- Medievalist
- Medieval philology
- Medieval philology
- Modern and medieval languages: focus on medieval historiography.
- Digital Literary Studies with a focus on French and Spanish literature.
- Digital Humanities with some more profound knowledge in History and Philosophy
- Lexicologist (medieval french), philologist, developer of dictionary writing systems
- I'm in Law, but my PhD was a diversion into Information Studies.
- I'm a linguist working between social sciences and the humanities.
- History
- Classical Archaeology
- I'm a rare books librarian so work with all subject areas.
- Classicist - Latin literature - focus on texts, textual analysis. Also interested in ancient philosophy, particularly Stoicism, and what happens to it in the Roman period. Plus I do classical reception.
- French literature (17th century) and digital humanities
- Literary studies & history of ideas
- Former senior TV professional, BA in education, now on a masters course MSc Digital Education.
- Digital media studies
- Classicist, but with a passion for modern English and German Literature
- Late Medieval and Early Modern Literature
- I'm a medievalist researcher.
- Classics, Latin literature
- Classical Philology
- Corpus Linguistics, Discourse Analysis and Translation Studies
- I'm an Arabist, but currently work in a library
- Digital Classicist
- Digital Libraries
- Classics, Ancient Greek and Latin, Archaeology, History
- history of ideas, in particular philosophy and language sciences, Eastern and Central Europe
- I work on French philology but currently I work more on manuscripts (codicology)
- Diplomatie et paléographie latine
- I am specialist in Classical and Medieval Latin, and work on History of Scientific (better: Medical) Texts
- I'm a Computing Scientist with a background in Ancient Greek.
- Computational Linguistics for Literature
- French literary history
- Computational linguistics applied to humanities data
- Classically trained in philology: research focused on late antique, medieval, and Byzantine texts in Greek, Syriac, and Latin.
- Medieval Historian
- I'm a PhD student in English philology, with a focus on manuscript studies and English historical linguistics.
- Medieval history of law
- epigraphy
Agriculture

My background is in Romance Philology and genetic philology. I researched on digital philology but currently I'm interested in how technology affects learning and teaching.

Manuscript Studies

Romance Philology

Codicologist and palaeographer specialising in Anglo-Saxon Manuscript culture now working in DH/archives/information studies.

I studied library and information science and Romance languages and am currently managing Digital Humanities Projects.

Medieval Medicine

Classicism, current focus is on Latin epigraphy (Middle Ages and Early Modern Period)

I am researching the intersection of art and science.

Biblical Studies (New Testament, Textual Criticism, Greek)

I'm a medievalist interested in textual criticism and digital tools for working in Spanish manuscripts and early printed books

My focus is on library and information science, digital libraries, digital humanities, digital cultural heritage

Late medieval-early modern England—approx. 1400-1550.

greek law, greek epigraphy (ancient)

I'm a Classicist, primarily a philologist, with interests in Greek religion and the reception of Greek literature by early Christians.

I'm a digital humanist in an English department who focuses on digital literature and adaptations.

Languages and literature

I am a citizen scientist involved in a variety of digital humanity and natural history crowdsourcing projects and am also interested in crowdsourcing as an academic discipline although I am not affiliated with any academic institution.

I'm a Classical Philologist.

I have a background in Social Sciences but I am currently focusing on History

Classicism, focus on classical archaeology

Sociology

Shakespeare Studies

Digital Humanities/Digital Archiving

I'm an English and Scottish Language and Literature but also work across broader Humanities subjects

Generally work in English and Scottish Language and Literature but also work across broader Humanities subjects

I am a modernist

Classicism specialised Roman Law and Ancient History

Modern Literatures

My subject area is the liturgical medieval chant

Medieval History/Art History

digital philology

digital philology

I'm a digital librarian with a background in English literature.

I'm a philosopher with responsibilities in digital editing.

British Literature

Area Studies, Early Modern Printing, East-Asia

Art History / Digital Art History

Jewish Studies - Rabbinics
Classicist with a focus on Classical Philology.

I'm a Classicist and now work in publishing
History of daily life
Linguistics
eighteenth century English literature and culture
Digital History
I am an engineer, my main background is on automation. Worked mainly in the power system and the cyber security sector. Recently got an MS degree on classical letters, with a thesis on a digital edition of Archilocus.

Classicist / medievalist
I'm a Classicist with a focus on Ancient Greek Literature.

More on Folklore and Digital Archiving
I am working on folklore and Digital Archiving

I'm linguistically oriented classicist using all kinds of sources, text-object-visual

I am interested particularly in Epistemology and Computer Sciences because I feel that both fields are interconnected in many ways

I find many topics connected in both disciplines. Computer Sciences is a privileged part of our culture and Philosophy is originated in and compromised with the culture.

Ancient History, Military History
I am a scholar of modern and contemporanea Italian Literature, and my focus for my PhD research is now about Italian Literature connection to the Digital Humanities

Surgery, History of Medicine, History in general
Scholar of modern and contemporary Italian Literature, which I want to study by applying Digital Humanities

Surgery, History of Medicine, History in general
Surgery, History of Medicine, History in general
I am a historian focusing on all aspects of the sixteenth century

historical metrology
Digital Humanities, creating digital editions

medieval legal historian
I'm a Sinologist, my main work consists of creating and maintaining digital texts and related tools

I'm a historian of the ancient Mediterranean with philological and epigraphical training

Text in the digital age
I'm working in the field of Patristics

Textual critic
I am a Latinist who also has an interest in digital editing.

I'm a medievalist (late medieval-early modern)

I'm a Classicist but my current focus is on Classical Philology

Theatre director specializing in Early Modern English drama. One-time professor of theatrical literature. Lecturer and speaker focusing on Early Modern English drama with a focus on Shakespeare and his contemporaries.

Classicist focused on language pedagogy.

Church historian
Digital Humanities, combining Phd in English with a career in IT.

Medievalist by training (manuscripts and historiography), with substantial work experience in humanities computing / digital humanities (textual criticism, codicology, American literature). Now an independent scholar in medieval MSS again while newly employed in the software industry. I can't tick "non-academic professional" after so many years with universities. :)

Particular interests in linguistics
Classicist working in medieval philosophy
Composition and Rhetoric, and Visual Rhetoric

History

History
I am an historian of the transmission of texts and sources in the Middle Ages

architectural historian and musicologist developing database-driven web-portals; focus is on long-time persistence of digital data (mainly, but not restricted to) in the humanities

Renaissance Italian literature and culture
Ancient (Roman) history
I am an English medievalist, focusing on philology, textual criticism and editing.
Mostly library and information studies, which tends to branch out into a lot of different subject areas depending on the research
I am focussed on Military History, primarily regarding both World Wars, and ego documents (letters and correspondences, diaries, autobiographies).
Historian (Diplomatics, paleography), now encharged in a digital humanities position
Ancient Near Eastern Philology / Hittitology
English Language and Applied Linguistics
Corpus stylistics, literary linguistics, historical sociolinguistics, and digital humanities
Long c18th History
Renaissance studies, mainly renaissance intellectual history/philosophy
Literary historian currently working on theatre history
Medieval manuscripts and book history, interested in physicality over text.
I'm an early modern historian focusing on late sixteenth/early seventeenth century cheap print.
History of the book, with focus on materiality and commercial aspects of printing & booktrade
Early modern history and literature
Early modern english scholar
I teach British literature before 1800 as well as a variety of writing classes. My research focuses on accessibility and usability in digital environments.
English Literature; Children's Literature; Drama
Food History
English Renaissance Drama
medieval history, focus on editing medieval latin texts
I'm a historian by training but my research is specifically related to digital scholarly editions - so I generally say my subject area is digital humanities
I'm an IS professional working in cultural heritage digitization and textual scholarship
Hispanic philologist but working on digital humanities
Literary scholar with a focus on digital scholarly editions
Mediaevalist, Jewish studies
I'm a philologist, though my current focus includes education and modern languages education.
Early modern history
Early American literature
My focus is on Spatial History and the History of Vulcanology and Seismology.
English Literature
Digital Humanities
Ciberanthropologist
I am trained in Editorship of Polish Literature and my dissertation was in 16th century translation of the Book of Psalms, but I am more interested in meta research about editorial studies rather than in particular literary epoque. Lately I've been studying User Research.
Corpus and Computational Linguistics in Greek
I'm a historian using and editing medieval sources to analyse medieval society
Digital Librarian
Philosophy
Medieval literatures and cultures
I'm a medievalist who focuses on textual criticism and translation in late medieval England.
Latin médiéval, Histoire médiévale
Literature scholar with focus on contemporary experimental literature
Classicist
I work on all facets of early modern letter-writing, from linguistic to material.
Journalism, history, new media
Medieval English Literature
19th century periodicals and legal history.

Archaeology

Early Medieval language literature and culture

Medieval historian (women and religion)

Biblical studies

My research focuses on linguistic and rhetorical features of ancient copies and translations of the Hebrew Bible

Early American Literature

research infrastructure (especially for DH)

Classicist doing archaeology and history.

I'm an archaeologist but currently working on legacy data in museum collections

Professional Writing, with a bend toward the technical: IA, procedures, instructional text

Medievalist who has a deep background (previous career) in programming and databases

I am a digital archaeologist focusing on 3D visualisation in prehistory

data-driven sociologist of culture and science

Textual Scholar, Digital Humanities & Caribbean Literature/Culture

Classics, specifically Roman history and religion

Computer scientist focused on interaction design with a background in DH related to textual criticism (Donne, Cervantes, New Testament and a few others).

Digital art history

Classical Philologist: Late Antiquity and Byzantine poetry. Philology. Text crit

Classicalist working at the moment on Classical Philology

I'm a Classicist working at the moment on Classical Philology

I'm a Classicist working currently on Classical Philology

Musicology

Classical Philologist: mostly textual criticism, mostly Latin, mostly verse.

I'm an Italian philologist and I'm focused on documentary texts written in twentieth century, specifically on private letters composed during the first world war.

Social history of the late Ottoman Middle East, media history, conceptual history

Archivist & historian working as a professional archivist in a university research centre.

Library information science + DH + classics

Historian, with focus on interdisciplinary field known as "history of the book"

Specialist in material culture of the ancient Near East with a focus on demonology

Bureaucracy

I'm a philologist, primarily a Celticist, but I branch out into Old English and Latin too.

Originally Art Historian, I take care of technical aspects and data modelling in multiple digital textual editions while maintaining data and web presentations

I take care of technology and data modelling of digital textual editions and also maintain data and web frontends

Art history, (A)H

Modern Comparative History

Classical Philology

Bureaucracy
### Question 4
What do you primarily seek out in a digital edition? Choose one of the following options:

<table>
<thead>
<tr>
<th>Choices</th>
<th>Absolute frequency</th>
<th>Cum. absolute frequency</th>
<th>Relative frequency</th>
<th>Cum. relative frequency</th>
<th>Adjusted relative frequency</th>
<th>Cum. adjusted relative frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acquisition of digitised versions of texts (and images, if available)</td>
<td>72</td>
<td>72</td>
<td>20.69%</td>
<td>20.69%</td>
<td>26.97%</td>
<td>26.97%</td>
</tr>
<tr>
<td>Acquisition of digitised versions of texts (and images, if available)</td>
<td>92</td>
<td>164</td>
<td>26.44%</td>
<td>47.13%</td>
<td>34.46%</td>
<td>61.42%</td>
</tr>
<tr>
<td>A complete educational resource, including interactivity and aesthetics</td>
<td>85</td>
<td>249</td>
<td>24.43%</td>
<td>71.55%</td>
<td>31.84%</td>
<td>93.26%</td>
</tr>
<tr>
<td>Other (please specify)</td>
<td>18</td>
<td>267</td>
<td>5.17%</td>
<td>76.72%</td>
<td>6.74%</td>
<td>100%</td>
</tr>
</tbody>
</table>

**Sum:** 267

**Not answered:** 81

**Total answered:** 267

**Average:** 2.18  **Minimum:** 1  **Variance:** 0.83  **Median:** 2  **Maximum:** 4  **Std. deviation:** 0.91

---

**Last choice text input**

- Efficient usability (reading, storing, citing, searching etc)
- Information
- Not sure any of these definitions work. I seek digital texts (and images) to inform research, which later may be published.
- All three answers
- All of the above.
- All

**A reconstruction of the lost original of a text starting from the survived copies of its and making available the materials (witnesses, transcriptions), textual analysis tools and, if possible, automatised management of the variants**

**Possibility to search the text**

- It depends
- All of the above

**A good critical text with an apparatus**

**Ability to make annotations and possibly share them. Acquisition is also desired, but ideally in a version-controlled form.**

**All of the above**

**The private/public distinction isn't clear. I access publically re-usable material for private use. I also publish some publically available material. 'Acquisition' seems to pertain more to material things not access/reading on the internet.**
I don't understand the distinctions made above
Question 5
How important is the scholarly component of a digital edition for you?

[Scholarly = “A reproduction of documents without critical examination is not scholarly editing. A facsimile is not a scholarly edition.”

(Patrick Sahle, see http://www.digitale-edition.de/vlet-about.html]

---

**Frequency table**

<table>
<thead>
<tr>
<th>Levels</th>
<th>Absolute frequency</th>
<th>Cum. absolute frequency</th>
<th>Relative frequency</th>
<th>Cum. relative frequency</th>
<th>Adjusted relative frequency</th>
<th>Cum. adjusted relative frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>No opinion</td>
<td>4</td>
<td>4</td>
<td>1.15%</td>
<td>1.15%</td>
<td>1.7%</td>
<td>1.7%</td>
</tr>
<tr>
<td>1 (Not at all important)</td>
<td>3</td>
<td>7</td>
<td>0.86%</td>
<td>2.01%</td>
<td>1.28%</td>
<td>2.98%</td>
</tr>
<tr>
<td>2</td>
<td>16</td>
<td>23</td>
<td>4.6%</td>
<td>6.61%</td>
<td>6.81%</td>
<td>9.79%</td>
</tr>
<tr>
<td>3</td>
<td>33</td>
<td>56</td>
<td>9.48%</td>
<td>16.09%</td>
<td>14.04%</td>
<td>23.83%</td>
</tr>
<tr>
<td>4</td>
<td>65</td>
<td>121</td>
<td>18.68%</td>
<td>34.77%</td>
<td>27.66%</td>
<td>51.49%</td>
</tr>
<tr>
<td>5 (Very important)</td>
<td>114</td>
<td>235</td>
<td>32.76%</td>
<td>67.53%</td>
<td>48.51%</td>
<td>100%</td>
</tr>
<tr>
<td><strong>Sum:</strong></td>
<td><strong>235</strong></td>
<td><strong>-</strong></td>
<td><strong>67.53%</strong></td>
<td><strong>-</strong></td>
<td><strong>100%</strong></td>
<td><strong>-</strong></td>
</tr>
<tr>
<td>Not answered:</td>
<td>113</td>
<td>-</td>
<td>32.47%</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Average:</td>
<td>4.17</td>
<td>Minimum: 1 Variance: 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median:</td>
<td>4</td>
<td>Maximum: 5 Std. deviation: 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Total answered: 235
Question 6
How important is knowing the production life-span or duration of a digital edition (begin and end years)?

<table>
<thead>
<tr>
<th>Levels</th>
<th>Absolute frequency</th>
<th>Cum. absolute frequency</th>
<th>Relative frequency</th>
<th>Cum. relative frequency</th>
<th>Adjusted relative frequency</th>
<th>Cum. adjusted relative frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>No opinion</td>
<td>16</td>
<td>16</td>
<td>4.6%</td>
<td>4.6%</td>
<td>6.81%</td>
<td>6.81%</td>
</tr>
<tr>
<td>1 (Not at all important)</td>
<td>14</td>
<td>30</td>
<td>4.02%</td>
<td>8.62%</td>
<td>5.96%</td>
<td>12.77%</td>
</tr>
<tr>
<td>2</td>
<td>37</td>
<td>67</td>
<td>10.63%</td>
<td>19.25%</td>
<td>15.74%</td>
<td>28.51%</td>
</tr>
<tr>
<td>3</td>
<td>42</td>
<td>109</td>
<td>12.07%</td>
<td>31.32%</td>
<td>17.87%</td>
<td>46.38%</td>
</tr>
<tr>
<td>4</td>
<td>69</td>
<td>178</td>
<td>19.83%</td>
<td>51.15%</td>
<td>29.36%</td>
<td>75.74%</td>
</tr>
<tr>
<td>5 (Very important)</td>
<td>57</td>
<td>235</td>
<td>16.38%</td>
<td>67.53%</td>
<td>24.26%</td>
<td>100%</td>
</tr>
<tr>
<td>Sum:</td>
<td>235</td>
<td>-</td>
<td>67.53%</td>
<td>-</td>
<td>100%</td>
<td>-</td>
</tr>
<tr>
<td>Not answered:</td>
<td>113</td>
<td>-</td>
<td>32.47%</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Average:</td>
<td>3.54</td>
<td>Minimum: 1</td>
<td>Variance: 1.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median:</td>
<td>4</td>
<td>Maximum: 5</td>
<td>Std. deviation: 1.22</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Total answered: 235
**Question 7**

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

![Bar chart showing responses to the question](chart.png)

<table>
<thead>
<tr>
<th>Levels</th>
<th>Absolute frequency</th>
<th>Cum. absolute frequency</th>
<th>Relative frequency</th>
<th>Cum. relative frequency</th>
<th>Adjusted relative frequency</th>
<th>Cum. adjusted relative frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>No opinion</td>
<td>7</td>
<td>7</td>
<td>2.01%</td>
<td>2.01%</td>
<td>2.98%</td>
<td>2.98%</td>
</tr>
<tr>
<td>1 (Not at all important)</td>
<td>16</td>
<td>23</td>
<td>4.6%</td>
<td>6.61%</td>
<td>6.81%</td>
<td>9.79%</td>
</tr>
<tr>
<td>2</td>
<td>30</td>
<td>53</td>
<td>8.62%</td>
<td>15.23%</td>
<td>12.77%</td>
<td>22.55%</td>
</tr>
<tr>
<td>3</td>
<td>70</td>
<td>123</td>
<td>20.11%</td>
<td>35.34%</td>
<td>29.79%</td>
<td>52.34%</td>
</tr>
<tr>
<td>4</td>
<td>57</td>
<td>180</td>
<td>16.38%</td>
<td>51.72%</td>
<td>24.26%</td>
<td>76.6%</td>
</tr>
<tr>
<td>5 (Very important)</td>
<td>55</td>
<td>235</td>
<td>15.8%</td>
<td>67.53%</td>
<td>23.4%</td>
<td>100%</td>
</tr>
<tr>
<td><strong>Sum:</strong></td>
<td><strong>235</strong></td>
<td></td>
<td><strong>67.53%</strong></td>
<td></td>
<td></td>
<td><strong>100%</strong></td>
</tr>
<tr>
<td>Not answered</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total answered:</strong></td>
<td><strong>235</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **Average:** 3.46
- **Minimum:** 1
- **Variance:** 1.42
- **Median:** 3
- **Maximum:** 5
- **Std. deviation:** 1.19

- **Total answered:** 235
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.)?

<table>
<thead>
<tr>
<th>Levels</th>
<th>Absolute frequency</th>
<th>Cum. absolute frequency</th>
<th>Relative frequency</th>
<th>Cum. relative frequency</th>
<th>Adjusted relative frequency</th>
<th>Cum. adjusted relative frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (Not at all important)</td>
<td>6</td>
<td>6</td>
<td>1.72%</td>
<td>1.72%</td>
<td>2.55%</td>
<td>2.55%</td>
</tr>
<tr>
<td>2</td>
<td>19</td>
<td>25</td>
<td>5.46%</td>
<td>7.18%</td>
<td>8.09%</td>
<td>10.64%</td>
</tr>
<tr>
<td>3</td>
<td>44</td>
<td>69</td>
<td>12.64%</td>
<td>19.83%</td>
<td>18.72%</td>
<td>29.36%</td>
</tr>
<tr>
<td>4</td>
<td>55</td>
<td>124</td>
<td>15.8%</td>
<td>35.63%</td>
<td>23.4%</td>
<td>52.77%</td>
</tr>
<tr>
<td>5 (Very important)</td>
<td>111</td>
<td>235</td>
<td>31.9%</td>
<td>67.53%</td>
<td>47.23%</td>
<td>100%</td>
</tr>
<tr>
<td>Sum:</td>
<td>235</td>
<td></td>
<td>67.53%</td>
<td>100%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not answered:</td>
<td>113</td>
<td></td>
<td>32.47%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average:</td>
<td>4.05</td>
<td>Minimum:</td>
<td>1</td>
<td>Variance:</td>
<td>1.22</td>
<td></td>
</tr>
<tr>
<td>Median:</td>
<td>4</td>
<td>Maximum:</td>
<td>5</td>
<td>Std. deviation:</td>
<td>1.1</td>
<td></td>
</tr>
</tbody>
</table>

Total answered: 235
**Question 9**

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

<table>
<thead>
<tr>
<th>Levels</th>
<th>Absolute frequency</th>
<th>Cum. absolute frequency</th>
<th>Relative frequency</th>
<th>Cum. relative frequency</th>
<th>Adjusted relative frequency</th>
<th>Cum. adjusted relative frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>No opinion</td>
<td>1</td>
<td>1</td>
<td>0.29%</td>
<td>0.29%</td>
<td>0.43%</td>
<td>0.43%</td>
</tr>
<tr>
<td>1 (Not at all important)</td>
<td>3</td>
<td>4</td>
<td>0.86%</td>
<td>1.15%</td>
<td>1.28%</td>
<td>1.7%</td>
</tr>
<tr>
<td>2</td>
<td>9</td>
<td>13</td>
<td>2.59%</td>
<td>3.74%</td>
<td>3.83%</td>
<td>5.53%</td>
</tr>
<tr>
<td>3</td>
<td>48</td>
<td>61</td>
<td>13.79%</td>
<td>17.53%</td>
<td>20.43%</td>
<td>25.96%</td>
</tr>
<tr>
<td>4</td>
<td>83</td>
<td>144</td>
<td>23.85%</td>
<td>41.38%</td>
<td>35.32%</td>
<td>61.28%</td>
</tr>
<tr>
<td>5 (Very important)</td>
<td>91</td>
<td>235</td>
<td>26.15%</td>
<td>67.53%</td>
<td>38.72%</td>
<td>100%</td>
</tr>
<tr>
<td><strong>Sum:</strong></td>
<td><strong>235</strong></td>
<td></td>
<td><strong>67.53%</strong></td>
<td><strong>100%</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Not answered:</strong></td>
<td><strong>113</strong></td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

**Average:** 4.07, **Minimum:** 1, **Variance:** 0.86, **Median:** 4, **Maximum:** 5, **Std. deviation:** 0.93

Total answered: 235
Question 10
How important is the provision of high quality digital images upon which the digital edition builds?

Frequency table

<table>
<thead>
<tr>
<th>Levels</th>
<th>Absolute frequency</th>
<th>Cum. absolute frequency</th>
<th>Relative frequency</th>
<th>Cum. relative frequency</th>
<th>Adjusted relative frequency</th>
<th>Cum. adjusted relative frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>No opinion</td>
<td>2</td>
<td>2</td>
<td>0.57%</td>
<td>0.57%</td>
<td>0.85%</td>
<td>0.85%</td>
</tr>
<tr>
<td>1 (Not at all important)</td>
<td>7</td>
<td>9</td>
<td>2.01%</td>
<td>2.59%</td>
<td>2.98%</td>
<td>3.83%</td>
</tr>
<tr>
<td>2</td>
<td>23</td>
<td>32</td>
<td>6.61%</td>
<td>9.2%</td>
<td>9.79%</td>
<td>13.62%</td>
</tr>
<tr>
<td>3</td>
<td>51</td>
<td>83</td>
<td>14.66%</td>
<td>23.85%</td>
<td>21.7%</td>
<td>35.32%</td>
</tr>
<tr>
<td>4</td>
<td>63</td>
<td>146</td>
<td>18.1%</td>
<td>41.95%</td>
<td>26.81%</td>
<td>62.13%</td>
</tr>
<tr>
<td>5 (Very important)</td>
<td>89</td>
<td>235</td>
<td>25.57%</td>
<td>67.53%</td>
<td>37.87%</td>
<td>100%</td>
</tr>
<tr>
<td>Sum:</td>
<td>235</td>
<td>-</td>
<td>67.53%</td>
<td>-</td>
<td>100%</td>
<td>-</td>
</tr>
<tr>
<td>Not answered:</td>
<td>113</td>
<td>32.47%</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Average:</td>
<td>3.88</td>
<td>Minimum: 1</td>
<td>Variance: 1.26</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median:</td>
<td>4</td>
<td>Maximum: 5</td>
<td>Std. deviation: 1.12</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Total answered: 235 }
Question 11
How important is the availability of advanced functionality and browsing, such as indices, filtering, searching and Application Programming Interfaces (API)?

<table>
<thead>
<tr>
<th>Levels</th>
<th>Absolute frequency</th>
<th>Cum. absolute frequency</th>
<th>Relative frequency</th>
<th>Cum. relative frequency</th>
<th>Adjusted relative frequency</th>
<th>Cum. adjusted relative frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (Not at all important)</td>
<td>4</td>
<td>4</td>
<td>1.15%</td>
<td>1.15%</td>
<td>1.7%</td>
<td>1.7%</td>
</tr>
<tr>
<td>2</td>
<td>9</td>
<td>13</td>
<td>2.59%</td>
<td>3.74%</td>
<td>3.83%</td>
<td>5.53%</td>
</tr>
<tr>
<td>3</td>
<td>24</td>
<td>37</td>
<td>6.9%</td>
<td>10.63%</td>
<td>10.21%</td>
<td>15.74%</td>
</tr>
<tr>
<td>4</td>
<td>80</td>
<td>117</td>
<td>22.99%</td>
<td>33.91%</td>
<td>34.04%</td>
<td>49.79%</td>
</tr>
<tr>
<td>5 (Very important)</td>
<td>118</td>
<td>235</td>
<td>33.91%</td>
<td>67.53%</td>
<td>50.21%</td>
<td>100%</td>
</tr>
<tr>
<td>Sum:</td>
<td>235</td>
<td>-</td>
<td>67.53%</td>
<td>-</td>
<td>100%</td>
<td>-</td>
</tr>
<tr>
<td>Not answered:</td>
<td>113</td>
<td>-</td>
<td>32.47%</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Average:</td>
<td>4.27</td>
<td>Minimum:</td>
<td>1</td>
<td>Variance:</td>
<td>0.84</td>
<td></td>
</tr>
<tr>
<td>Median:</td>
<td>5</td>
<td>Maximum:</td>
<td>5</td>
<td>Std. deviation:</td>
<td>0.92</td>
<td></td>
</tr>
</tbody>
</table>

Total answered: 235
**Question 12**
How important is the possibility of consulting a digital edition's website in multiple languages other than English?

![Bar chart showing frequency distribution](chart.png)

### Frequency table

<table>
<thead>
<tr>
<th>Levels</th>
<th>Absolute frequency</th>
<th>Cum. absolute frequency</th>
<th>Relative frequency</th>
<th>Cum. relative frequency</th>
<th>Adjusted relative frequency</th>
<th>Cum. adjusted relative frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>No opinion</td>
<td>9</td>
<td>9</td>
<td>2.59%</td>
<td>2.59%</td>
<td>3.83%</td>
<td>3.83%</td>
</tr>
<tr>
<td>1 (Not at all important)</td>
<td>57</td>
<td>66</td>
<td>16.38%</td>
<td>18.97%</td>
<td>24.26%</td>
<td>28.09%</td>
</tr>
<tr>
<td>2</td>
<td>54</td>
<td>120</td>
<td>15.52%</td>
<td>34.48%</td>
<td>22.98%</td>
<td>51.06%</td>
</tr>
<tr>
<td>3</td>
<td>46</td>
<td>166</td>
<td>13.22%</td>
<td>47.7%</td>
<td>19.57%</td>
<td>70.64%</td>
</tr>
<tr>
<td>4</td>
<td>43</td>
<td>209</td>
<td>12.36%</td>
<td>60.06%</td>
<td>18.3%</td>
<td>88.94%</td>
</tr>
<tr>
<td>5 (Very important)</td>
<td>26</td>
<td>235</td>
<td>7.47%</td>
<td>67.53%</td>
<td>11.06%</td>
<td>100%</td>
</tr>
<tr>
<td>Sum:</td>
<td>235</td>
<td>-</td>
<td>67.53%</td>
<td>-</td>
<td>100%</td>
<td>-</td>
</tr>
<tr>
<td>Not answered:</td>
<td>113</td>
<td>-</td>
<td>32.47%</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Average: 2.68  Minimum: 1  Variance: 1.8  Median: 3  Maximum: 5  Std. deviation: 1.34

Total answered: 235
Question 13
How important is the provision of data in Open Source/Access formats?

<table>
<thead>
<tr>
<th>Levels</th>
<th>Absolute frequency</th>
<th>Cum. absolute frequency</th>
<th>Relative frequency</th>
<th>Cum. relative frequency</th>
<th>Adjusted relative frequency</th>
<th>Cum. adjusted relative frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>No opinion</td>
<td>8</td>
<td>8</td>
<td>2.3%</td>
<td>2.3%</td>
<td>3.4%</td>
<td>3.4%</td>
</tr>
<tr>
<td>1 (Not at all important)</td>
<td>3</td>
<td>11</td>
<td>0.86%</td>
<td>3.16%</td>
<td>1.28%</td>
<td>4.68%</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>17</td>
<td>1.72%</td>
<td>4.89%</td>
<td>2.55%</td>
<td>7.23%</td>
</tr>
<tr>
<td>3</td>
<td>37</td>
<td>54</td>
<td>10.63%</td>
<td>15.52%</td>
<td>15.74%</td>
<td>22.98%</td>
</tr>
<tr>
<td>4</td>
<td>51</td>
<td>105</td>
<td>14.66%</td>
<td>30.17%</td>
<td>21.7%</td>
<td>44.68%</td>
</tr>
<tr>
<td>5 (Very important)</td>
<td>130</td>
<td>235</td>
<td>37.36%</td>
<td>67.53%</td>
<td>55.32%</td>
<td>100%</td>
</tr>
<tr>
<td>Sum:</td>
<td>235</td>
<td>-</td>
<td>67.53%</td>
<td>-</td>
<td>100%</td>
<td>-</td>
</tr>
<tr>
<td>Not answered:</td>
<td>113</td>
<td>-</td>
<td>32.47%</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Average:</td>
<td>4.32</td>
<td>Minimum: 1</td>
<td>Variance: 0.86</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median:</td>
<td>5</td>
<td>Maximum: 5</td>
<td>Std. deviation: 0.93</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Total answered: 235
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)?

<table>
<thead>
<tr>
<th>Levels</th>
<th>Absolute frequency</th>
<th>Cum. absolute frequency</th>
<th>Relative frequency</th>
<th>Cum. relative frequency</th>
<th>Adjusted relative frequency</th>
<th>Cum. adjusted relative frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>No opinion</td>
<td>9</td>
<td>9</td>
<td>2.59%</td>
<td>2.59%</td>
<td>3.83%</td>
<td>3.83%</td>
</tr>
<tr>
<td>1 (Not at all important)</td>
<td>27</td>
<td>36</td>
<td>7.76%</td>
<td>10.34%</td>
<td>11.49%</td>
<td>15.32%</td>
</tr>
<tr>
<td>2</td>
<td>56</td>
<td>92</td>
<td>16.09%</td>
<td>26.44%</td>
<td>23.83%</td>
<td>39.15%</td>
</tr>
<tr>
<td>3</td>
<td>75</td>
<td>167</td>
<td>21.55%</td>
<td>47.99%</td>
<td>31.91%</td>
<td>71.06%</td>
</tr>
<tr>
<td>4</td>
<td>43</td>
<td>210</td>
<td>12.36%</td>
<td>60.34%</td>
<td>18.3%</td>
<td>89.36%</td>
</tr>
<tr>
<td>5 (Very important)</td>
<td>25</td>
<td>235</td>
<td>7.18%</td>
<td>67.53%</td>
<td>10.64%</td>
<td>100%</td>
</tr>
<tr>
<td>Sum:</td>
<td>235</td>
<td></td>
<td>67.53%</td>
<td>100%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not answered:</td>
<td>113</td>
<td></td>
<td>32.47%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average:</td>
<td>2.92</td>
<td>Minimum:</td>
<td>1</td>
<td>Variance:</td>
<td>1.36</td>
<td></td>
</tr>
<tr>
<td>Median:</td>
<td>3</td>
<td>Maximum:</td>
<td>5</td>
<td>Std. deviation:</td>
<td>1.17</td>
<td></td>
</tr>
</tbody>
</table>

Total answered: 235
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)?

<table>
<thead>
<tr>
<th>Levels</th>
<th>Absolute frequency</th>
<th>Cum. absolute frequency</th>
<th>Relative frequency</th>
<th>Cum. relative frequency</th>
<th>Adjusted relative frequency</th>
<th>Cum. adjusted relative frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>No opinion</td>
<td>5</td>
<td>5</td>
<td>1.44%</td>
<td>1.44%</td>
<td>2.13%</td>
<td>2.13%</td>
</tr>
<tr>
<td>1 (Not at all important)</td>
<td>37</td>
<td>42</td>
<td>10.63%</td>
<td>12.07%</td>
<td>15.74%</td>
<td>17.87%</td>
</tr>
<tr>
<td>2</td>
<td>54</td>
<td>96</td>
<td>15.52%</td>
<td>27.59%</td>
<td>22.98%</td>
<td>40.85%</td>
</tr>
<tr>
<td>3</td>
<td>47</td>
<td>143</td>
<td>13.51%</td>
<td>41.09%</td>
<td>20%</td>
<td>60.85%</td>
</tr>
<tr>
<td>4</td>
<td>54</td>
<td>197</td>
<td>15.52%</td>
<td>56.61%</td>
<td>22.98%</td>
<td>83.83%</td>
</tr>
<tr>
<td>5 (Very important)</td>
<td>38</td>
<td>235</td>
<td>10.92%</td>
<td>67.53%</td>
<td>16.17%</td>
<td>100%</td>
</tr>
<tr>
<td>Sum</td>
<td>235</td>
<td>-</td>
<td>67.53%</td>
<td>-</td>
<td>100%</td>
<td>-</td>
</tr>
<tr>
<td>Not answered</td>
<td>113</td>
<td>-</td>
<td>32.47%</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Average</td>
<td>3.01</td>
<td>Minimum</td>
<td>1</td>
<td>Variance</td>
<td>1.78</td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>3</td>
<td>Maximum</td>
<td>5</td>
<td>Std. deviation</td>
<td>1.33</td>
<td></td>
</tr>
<tr>
<td>Total answered</td>
<td>235</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**Question 16**
How important is the possibility of downloading and reusing the data published within a digital edition?

![Bar chart showing frequency distribution of responses to Question 16](chart.png)

<table>
<thead>
<tr>
<th>Levels</th>
<th>Absolute frequency</th>
<th>Cum. absolute frequency</th>
<th>Relative frequency</th>
<th>Cum. relative frequency</th>
<th>Adjusted relative frequency</th>
<th>Cum. adjusted relative frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>No opinion</td>
<td>1</td>
<td>1</td>
<td>0.29%</td>
<td>0.29%</td>
<td>0.43%</td>
<td>0.43%</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>11</td>
<td>2.87%</td>
<td>3.16%</td>
<td>4.26%</td>
<td>4.68%</td>
</tr>
<tr>
<td>3</td>
<td>25</td>
<td>36</td>
<td>7.18%</td>
<td>10.34%</td>
<td>10.64%</td>
<td>15.32%</td>
</tr>
<tr>
<td>4</td>
<td>62</td>
<td>98</td>
<td>17.82%</td>
<td>28.16%</td>
<td>26.38%</td>
<td>41.7%</td>
</tr>
<tr>
<td>5 (Very important)</td>
<td>137</td>
<td>235</td>
<td>39.37%</td>
<td>67.53%</td>
<td>58.3%</td>
<td>100%</td>
</tr>
<tr>
<td><strong>Sum:</strong></td>
<td>235</td>
<td></td>
<td><strong>67.53%</strong></td>
<td><strong>100%</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Not answered:</strong></td>
<td>113</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Average:</strong></td>
<td>4.39</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Minimum:</strong></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Variance:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.71</td>
</tr>
<tr>
<td><strong>Maximum:</strong></td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.84</td>
</tr>
<tr>
<td><strong>Total answered:</strong></td>
<td>235</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

24 / 35
### Frequency table

<table>
<thead>
<tr>
<th>Choices</th>
<th>Absolute frequency</th>
<th>Cum. absolute frequency</th>
<th>Relative frequency by choice</th>
<th>Relative frequency</th>
<th>Cum. relative frequency</th>
<th>Adjusted relative frequency</th>
<th>Cum. adjusted relative frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teaching</td>
<td>170</td>
<td>170</td>
<td>31.6%</td>
<td>48.85%</td>
<td>48.85%</td>
<td>72.96%</td>
<td>72.96%</td>
</tr>
<tr>
<td>Textual analyses (e.g. morpho-syntactical parsing)</td>
<td>166</td>
<td>336</td>
<td>30.86%</td>
<td>47.7%</td>
<td>96.55%</td>
<td>71.24%</td>
<td>144.21%</td>
</tr>
<tr>
<td>Imaging analyses (e.g. computer vision)</td>
<td>50</td>
<td>386</td>
<td>9.29%</td>
<td>14.37%</td>
<td>110.92%</td>
<td>21.46%</td>
<td>165.67%</td>
</tr>
<tr>
<td>Corpus aggregation</td>
<td>115</td>
<td>501</td>
<td>21.38%</td>
<td>33.05%</td>
<td>143.97%</td>
<td>49.36%</td>
<td>215.02%</td>
</tr>
<tr>
<td>Other (please specify)</td>
<td>37</td>
<td>538</td>
<td>6.88%</td>
<td>10.63%</td>
<td>154.6%</td>
<td>15.88%</td>
<td>230.9%</td>
</tr>
</tbody>
</table>

**Sum:** 538  -  100%  -  -  -  -  -  -

**Not answered:** 115  -  -  33.05%  -  -  -  -

**Average:** 2.41  **Minimum:** 1  **Variance:** 1.72

**Median:** 2  **Maximum:** 5  **Std. deviation:** 1.31

**Total answered:** 233

---

**Last choice text input**

Arguably you don’t know what use might emerge

Analyses in general

Consultation for my own research; publishing the text in an article/offering analysis of it.

Specific annotation and translation

Sharing in my own digital project

Evaluation of stemmatic reconstruction

Research

Referencing, annotation

Linked open data

Search text

“Data” exclusive of historical, literary, and other values? I don’t generally need text as data.

My main use is of images and their associated data in a digital edition for uploading into Wikicommons for use in Wikipedia articles.

Full analysis as a set

Research in many ways

Reading

Research

Quotation

Research
Creating new intersecting 'editions'

Enjoyment and appreciation of the text.
Research of sources and parallels of intertextuality
Collect contextualised quotations for further use
Image reuse of various sorts, but I don't do computer vision
Traditional humanities research activities
Literary analyses

General uses, non academic uses
Textual analysis of the old fashioned close reading kind
with proper meta-tagging: network and timeline analysis
Re-editing

Data visualisation
research (I have no clue what you mean by the three above options)
My own research
Literary analysis.

Quote
### Frequency Table

<table>
<thead>
<tr>
<th>Choices</th>
<th>Absolute Frequency</th>
<th>Cumulative Absolute Frequency</th>
<th>Relative Frequency by Choice</th>
<th>Cumulative Relative Frequency</th>
<th>Adjusted Relative Frequency</th>
<th>Cumulative Adjusted Relative Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>XML</td>
<td>103</td>
<td>103</td>
<td>13.31%</td>
<td>29.6%</td>
<td>44.21%</td>
<td>44.21%</td>
</tr>
<tr>
<td>XML TEI</td>
<td>150</td>
<td>253</td>
<td>19.38%</td>
<td>43.1%</td>
<td>64.38%</td>
<td>108.58%</td>
</tr>
<tr>
<td>Plain Text (.txt)</td>
<td>124</td>
<td>377</td>
<td>16.02%</td>
<td>35.63%</td>
<td>108.33%</td>
<td>161.8%</td>
</tr>
<tr>
<td>PDF</td>
<td>122</td>
<td>499</td>
<td>15.76%</td>
<td>35.06%</td>
<td>143.39%</td>
<td>214.16%</td>
</tr>
<tr>
<td>ePub</td>
<td>35</td>
<td>534</td>
<td>4.52%</td>
<td>10.06%</td>
<td>153.45%</td>
<td>229.18%</td>
</tr>
<tr>
<td>Images (TIFF)</td>
<td>89</td>
<td>623</td>
<td>11.5%</td>
<td>25.57%</td>
<td>179.02%</td>
<td>267.38%</td>
</tr>
<tr>
<td>Images (PNG, JPG, JPEG)</td>
<td>135</td>
<td>758</td>
<td>17.44%</td>
<td>38.79%</td>
<td>217.82%</td>
<td>325.32%</td>
</tr>
<tr>
<td>Other (please specify)</td>
<td>16</td>
<td>774</td>
<td>2.07%</td>
<td>4.6%</td>
<td>222.41%</td>
<td>332.19%</td>
</tr>
<tr>
<td>Sum:</td>
<td>774</td>
<td>-</td>
<td>100%</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Not answered:</td>
<td>115</td>
<td>-</td>
<td>33.05%</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Average:</td>
<td>3.93</td>
<td>Minimum:</td>
<td>1</td>
<td>Variance:</td>
<td>4.54</td>
<td></td>
</tr>
<tr>
<td>Median:</td>
<td>4</td>
<td>Maximum:</td>
<td>8</td>
<td>Std. deviation:</td>
<td>2.13</td>
<td></td>
</tr>
</tbody>
</table>

**Total answered:** 233

---

Last choice text input:
- docx (true, believe it or not)
- Not JP(E)G, it's a lossy format!
- it depends
- json
- The question should start with "Which"
- I don't know.
- Docx
- Web APIs, web services, open standards, removal of access restrictions.
- SQL
- mobile app, web app
- HTML5
- Anything supported by pandoc
- don't care
- Csv
- Word
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?

Text input

My own digital edition: daa.aau.at
probably http://iospe.kcl.ac.uk/corpora/Byzantine/index.html, as being the most recent instance of EpiDoc; but anything accessible is welcome, at this stage.

no comment
Université de Caen : double encoding of critical apparatus : XML-TEI and notes dynamic interface scientific liability provided by the critical edition + images see also

Froissart Online Exhaustivity
http://gams.uni-graz.at/context:srbas Good browsing/search; advanced functionalities (data basket); multiple export formats

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.

I get the most out of basic pdf docs imported into Notability for iPad. Straight and simple.

For teaching, a simple PDF is perfect. Students can view them under a wide variety of contexts, and they don’t require a persistent internet connection for use. For my own research, any format is fine as long as a plain text version is also made available. Most Archive org documents meet this criteria, but something like Debates in the Digital Humanities doesn’t. The open access version of Debates does, however, does have its complete source available through github, and the full, plain text of it can be extracted from those files.

http://www.janeausten.ac.uk/index.html
Not really
http://saint-denis.encyc.sorbonne.fr
http://vergil.classics.upenn.edu A strength of the Vergil Project is that it gives some info about textual variants, and you have the option to display those inline. There’s also a lot of other data visible on a single screen, including concordance date, Homeric correspondences and relevant parts of various commentaries. You can ‘hover’ over a word to view its vowel quantities.

http://www.sitzextase.de/digital-muqtabas/
The Digital Fragmenta Historiorum Graecorum meets my needs. I would like an edition made with its principles and including the critical apparatus, maybe done through the Perseids platform to avoid copyright infringements.

LOEB Classical Library. The edition is very basic, but it has the text, the translation, the notes and search options. Brilliant if you do not have immediate access to the physical books.

The digitizing program of the Vatican Library and of the Bayerische Staatsbibliothek
DC3 edition pioneered by Hugh Cayless, https://www.youtube.com/watch?v=n-pWJR33AyE
http://www.documentacatholicomaonnia.eu – While this leaves much to be desired in terms of quality of the texts provided, it is extremely useful because of its wide scope and the fact that it is based on a well known and widely available corpus.

Bess of Hardwick’s Letters is great for pictures, editorial principles clearly described, and diplomatic transcriptions. However, even there I don’t think you can download all XML files at once.

Oldbaily, criminocorpus…. Excellent information retrieval and statistical analysis. Download and reuse options could be better.

Epigraphic Database Roma Epigraphic Database Heidelberg Epigraphic Database Bari EAGLE
oxford scholarly editions online - searchable and easy to cite

The Electronic Beowulf http://ebeowulf.uky.edu/

Although not in the immediate field of my studies or even a little because of that and the fear that being too much concerned with those text types might entail too specific requirements, I chose here the vangoghletters instead, since they are [all the more for the time in which they have been created] a comprehensive archive, which gives contextual information without forcing you to view it. They provide the photos, transcriptions, sketches, translations and enable for instance citability (through providing line numbers and through being a completed project). Some more texttechnological applications and especially visualizations such as a network of the people involved with weighted edges or a lemmatization, a lexicon, some preprocessing and also some statistical analyses of the letters (even as simple as mean length, number of named entities, how many in which years etc.) are in my opinion still missing but may be most interesting for scientific users; in this sense the edition of the Divina Commedia offers a little more, but since it is commercial and I have no login, an assessment is rather hard. Here, Plaoul and codex sinaiticus are rather more functional

Welscher Gast Digital, http://digi.ub.uni-heidelberg.de/wgd/ Among many things: - fine grain information about the graphic system (allographs, abbreviations, punctuations), and a configurable interface; - downloadable sources in XML/TEI; - synoptic visualisation of the manuscript transcriptions; - CC BY SA licence; - DOI for citability; - digital facsimile.

e-codices: Excellent high quality images and good metadata. I can do the rest.

[I would cite my own digital edition here, so I will abstain from answering this question]
A digital edition, that has good functionality for my needs, must have: text, apparatus, textual analysis, witness (with images), commentary, bibliography, translation, data search, possibility of downloading.

There is none

I like lots of editions but at the moment particularly http://digi.ub.uni-heidelberg.de/wgd/; it has weaknesses (i.e. too much information everywhere, I feel confused) but I think the way they capture information in facsimiles and make them in integral part of the edition is somehow innovative (i.e. http://www.wgd.materiale-textkulturen.de/illustrationen/motiv.php?m=1).

http://www.beckettarchive.org/

When I was studying Classics I used the Perseus Collections of Greek and Roman Materials. I found it useful because it includes a wide range of texts, and, as well as the original Latin and Greek texts, it has many translations, so sometimes I could find a text that wasn't available in the library or wherever I was. Also, I could search for a word within a work to find parallels (if I wasn't in a library with access to the online Thesaurus Linguae Graecae).

The Shelley-Godwin Archive, although not for materials within my period, looks like the type of edition I would use, providing images, different transcription views, and downloadable XML. It also hybridizes the single-text edition with the archive, allowing the study of multiple objects within the same platform. Within my own period of the English Middle Ages, the new online version of the Electronic Beowulf would be the gold star.

Not satisfied with currently available options.

The Online Variorum of the Origin of Species. It shows a synoptic view of variation which helps to understand the timing of the changes. It is a pity that a mobile App has not yet been developed.

Perseus Digital Library.

Hm. The only example I can think of isn't public, since it's one I'm building myself (and thus is uniquely designed to cater to my needs). In my experience, Byzantinists aren't always enthusiastic about this kind of technological approach, so I'm the only one I know who's working on something of this nature. My main concerns in making this edition for myself were 1.) to document where and how this particular scribe diverges from the canonical version of a given scriptural commentary, and 2.) to clarify what exactly this manuscript contains, to better understand how/why the copyist put it together in this way. It's also quite handy to have something I can do corpus analysis with!

Perseus comes closest for teaching and some research. I am probably not fully aware of what is available.

Beowulf online by Kevin Kiernan Dnate's Comedia by Peter Robinson Skaldic Project

The Corpus Corporum is an extremely useful and a joy to use (especially the Patrologia Latina). But depending on your criteria it might not even qualify as a digital edition, since it's just a copy of the 19th c. Patrologia Latina.

http://www.mlats.uzh.ch/MLS/

None are really god. Google books is almost always bad.


The Exeter Book CD Rom emphasised links between the digital image/ms page and the critical edition/textual notes, and connected the two in ways I'd not seen done as well before (at that point, c.2006). However! CD Rom no longer works (links don't work between edition and image; in fact, edition data gone altogether) and images though decent quality have been manipulated/flattenated, don't have a colour strip, don't have measurements, etc. (Electronic Beowulf - great its online (finally) the design very clunky and V-images not brilliantly represented. Essentially I need good quality images and transcriptions/edited text that I can do stuff with (reuse/download etc); if they're linked, even better!)

The way that Bible Software (Accordance, Logos) has implemented digital original languages is quite useful. Has mobile access, lexicons, apparatuses, notes, etc. Very accessible and the platform is already there.

I can't

Oxford's Holinshed Project (http://www.cems.ox.ac.uk/holinshed/); diplomatic transcription to plain (not TEI) text; parallel texts of different editions; annotations; easy, reliable copy-and-paste.

IAph 2007: It contains the texts and all the necessary metadata

Most volumes contained in the Biodiversity Heritage Library provide good functionality for my needs. It would be helpful if more of the individual articles in some of the journals contained in the Biodiversity Heritage Library had DOI's attached to them. The would enable easier citations. But the fact I am able to download images and pages for reuse makes this platform the most useful for me in my work.

http://www.walkeraert.org/collections/publications citation, images, context, primary resources, multimedia

Edgar Allan Poe Society of Baltimore is more like a complete Poe resource; it's very technologically but very thorough and extremely reliable as far as texts go. I use it for my own research and for teaching.

I have yet to find a linguistically searchable corpus that really meets my needs.

Perseus Odyssey

Olive Schreiner Letters Online www.oliveschreiner.org

Internet Shakespeare Editions from the University of Victoria. It maintains very high standards of accuracy and gives its stuff away.


The Digital Vercelli site (http://vbd.humnet.unipi.it/beta2/) and the tool that powers it are very good. It's possible to view facsimiles and transcriptions side by side, or switch to a purely facsimile or transcription view. It has decent search facilities, a good facsimile zoom tool and just feels like a very well put together system. (Note that I have no involvement with this site or the tool).

Bibliotheca Iuris Antiqui (BIA)
Folger digital texts: reliable modern editions, in a variety of formats (e.g. txt, tei xml). These can be downloaded or searched on the online platform.

- http://wittgensteinrepository.org/: Offers access to facsimiles of the Wittgenstein Nachlass + manuscript descriptions + (increasingly) transcriptions. - http://wab.uib.no/transform/wab.php?modus=opsjoner: Offers user-steered access (incl. filtering and presentation options) to the same materials

I consider the Modern Journals Project a very good digital edition. Technically, it's not an edition, but an archive--but this is why I like it so much. Each text provided is clustered natively within a meaningful context (a particular medium during a particular artistic movement). This allows specialists to run comparative analyses or historical/material readings on these digital editions of particular magazines while allowing students or the general public simply to access the texts. What this means for me is that the digital edition can play two roles: providing open-access availability to the public and providing researchers with new avenues for research.

https://tei2016app.acdh.oeaw.ac.at/ provides an api to retrieve data, is open source (code of the application) and open data no, there is none that I know of

http://www.catullusonline.org/CatullusOnline/index.php meets my needs because it is openly accessible and of a very high quality in terms of scholarship and availability of materials.

In addition to the facilities provided by the Perseus Digital Library and Perseids, platforms like EVT and tools like the Classical Text Editor seem to provide most of the needed features, on a case by case basis. My MS thesis (available on request) provides a mock-up of the digital edition I have designed, in view of providing a support to teaching Ancient Greek literature to university students. This integrates repositories, tools and other facilities available from the above.

http://elec.enc.sorbonne.fr/chroniqueslatines/

Not now.

Free Courses On line (Coursera, Open University, Stanford, MIT, etc.)

Petrarchian Library (Petrarch's Rerum Vulgarium Fragmenta) because its great functionality (images; text; writing)

Just about anything published by Tanner-Ritchie. Totally free of scanning problems; material difficult to find in libraries in my part of the world. Aside from the photographed hands and illegible gutters often encountered, my greatest problem is 19th c works scanned at such a low resolution that the footnotes (in 6 pt type) are entirely illegible. The worst material is the Early English Books collection.

https://www.briefedition.alfred-escher.ch – clear editorial principals, good user interface

Something like the Early English Laws project is useful because it lets you put things side by side and also copy things in a plain text format to use in other forms/mark-up outside of the browser.

My work at www.kanripo.org and www.mandoku.org tries to build editions that fit the bill.

PHI Greek Inscriptions, though only for personal research since it lacks translations. It provides easy and quick access to primary source evidence that I can later look up in published but less accessible versions when necessary, but online allows me to formulate conclusions and engage in scholarly conversations in a timely fashion.

Dickinson Classical Commentaries (http://dcc.dickinson.edu) - a little more interactivity/customization would be appreciated, but it is currently good enough to use in place of a book for those students who are comfortable using an online text. Actually no, I haven't found any edition yet that is available both as ePub and website with an API, so that it allows me both to just read and enjoy it on a tablet (or even as print) or use it as teaching material, and that allows me to create interesting new entry points by reusing elements of it through its API (or download facilities). I'm sure there must exist some, but apparently not in my areas of interest.


The online Froisart

Icelandic Saga Map http://sagamap.hi.is/is/ - Complete corpus - Well defined corpus - Reliable corpus and curation - No feature overkill. It does what it does well. - Utilises IT to augment the texts with features that would not be possible in print and with enhance our ability to appreciate and understand the texts. In this case, finding places across all texts, linking places to a map and linking the map to locations in the corpus.

I'm not sure I have found one that meets my needs for teaching. I teach research writing and digital editions of textbooks could help students see the inquiry process from idea to publication. It's all too formulaic and simplistic at this point to be more helpful to me than the materials I have gathered through past student examples.

Corpus Corporum (Philipp Roelli) - because the corpus is extremely rich, but it does not have the functionalities I would expect for a digital CRITICAL edition (the quality of each works depends of the edition used to be put in the corpus).

Verhandlungen des Deutschen Reichstags und seiner Vorläufer (Session Reports of the German Reichstag and its Precursors, http://www.reichstagsprotokolle.de/index.html) - offering reliable and scholarly fully re-useable information as full text resource - allowing cross-search through various resources

http://gams.uni-graz.at/context:srbas?mode=projekt Jahrrechnungen der Stadt Basel You can interact with the material by


Norton Shakespeare online provides (some) performance information alongside fully-edited text, glosses, and other information; allows comparison of different editions/formats (so interactive and flexible, with good user interface)

No.
I currently use EEBO, which has mostly everything that I need, but transcription of text is occasionally unhelpful/unavailable. The ability to search via STC nos is invaluable, as is the ability to study the text closely. The variant spelling function is useful, as is the fuzzy function. I love the Folger Folger Digital Texts Shakespeare (and associated API) http://www.folgerdigitaltexts.org/, http://www.folgerdigitaltexts.org/api

TEAMS medieval texts: explains what the source is, gives basic commentary, is clear it's not the ne plus ultra scholarly edition but is free, online, handy for quick reference when libraries are closed or the library I belong to doesn't have the scholarly edition available.

I don't know of any that meet my teaching needs--computers in classrooms still not universal. But folger digital editions are pretty good.

Nothing immediately comes to mind.

Cambridge Ben Jonson: it has proper scholarly apparatus and accurate texts.

I use Open Source Shakespeare a lot. Though it has problems, the search functions, line numbers, and provenance make it a reliable and trusted online edition of the complete works. I also use the Shakespeare app for iOS, which functions more or less the same as OSS, except that it has more text (includes the poems and a couple more plays). I use these two sources most frequently when I'm away from my physical editions at home and need to check a reference, or to copy-paste a large passage for analysis in my work. Though I will always check those digital editions against my Arden or my Norton, being able to access them on the fly is super handy. I most often use them for the search function when I'm trying to find a particular passage or for language analysis (how often does X word get used, etc.). When preparing texts for performance, I will generally start with the OSS text and reformat/cut from there, but PlayShakespeare.com and the Folger digital editions are also good sources for that. MIT Shakespeare is terrible and should be stricken from the internet.


I cannot recall at the moment.

Briefedition Leopold Wilhelm

I've never had the pleasure to use an edition that fulfilled my ideal. The Becerro Galicano of San Millán: easy access to the source material and a range of search tools http://www.oxfordscholarlyeditions.com/ an example of a simple, clever, and efficient interface

When earlier in the survey the idea of user interactivity and a "complete" experience was introduced, I realized how rare that is. I like the text part of the Bayeux Tapestry edition http://www.bayeuxtapestry.org.uk/index.htm and the user experience of the Letter of Prester John (though it is missing some vital elements of an actual scholarly edition) http://scalar.usc.edu/works/prester-john/index. Something that combines these would be ideal.

van Gogh Letters, http://vangoghletters.org/vgl/ - good survey, quick orientation, smart navigation/hierarchisation of information - transparency (working principles, definitions etc.)

Not really: every edition is a compromise. Some are excellent at accessing the texts but don't provide images (e.g. http://www.livesandletters.uk/odl/bodley/bodley.html ); others may have more contextual information and indexing and images, but require more clicking to actually get to the texts and/or images (e.g. https://www.bessoighthardwick.org/ ). I think there's probably a trade-off between adding bells and whistles and making something readily accessible.

Beowulf Keenan and to a much lesser extent Bayeux Tapestry Foys ( because this did not use actual images ) both good intros to scholarship around the two texts for U/G s with no / little experience of studying MSS

Not really: it's hard to come by into an edition that provides every source material needed to replicate the result.

There are very few digital editions in my field. One I have used recently is Memorable Days: the Emilie Davis Diaries. Valuable because makes available a text previously hard to access, including to the manuscript. It has drawbacks too.

Abby Covert's _How to Make Sense of Any Mess_, (http://www.howtomakesenseofanymess.com) turned out to be an interesting case for me this past year: I've worked with the physical paperback, the ebook, and the website versions of this text, and it turned out that the ebook was the least useful of the three because of it's degraded random access.

I can't identify one yet, except perhaps my own work to be published in the next year. I struggle to find digital editions which really address social history concerns.

https://elotroalex.github.io/ed/ (Full disclosure: I built it)

The Perseus Project combines open access texts with an accurate parsing/dictionary feature.

Oxyrhyncus at Oxford

Dante Alighieri: Commedia. A Digital Edition ...if it was open access Some strong points: High-resolution colour digital images of manuscripts and easy Zooming Collation of variants linked to a map Go directly to any view of any line in any witnessed from a menu

Working on my own because nothing available does what I want them to do.

The online Froissart project https://www.hrionline.ac.uk/onlinefroissart/

hmmm, difficult choice, those of perseus/oppure and digiliblt are good approximations

Can't think of a specific document, but essentials include: - PAGE NUMBERS!!! (so many editions leave these out, which means I can't offer them for student use as I expect students to cite page references) - An accessible footnote/endnote system - Searchability - Ability to change size of fonts

31 / 35
<table>
<thead>
<tr>
<th>Yellow 90s Online (The Yellow Book)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
</tr>
<tr>
<td>none</td>
</tr>
<tr>
<td>Only partly. The Confessio edition is well done but interaction with the text is limited and it isn't downloadable. Granted it oriented towards the general public.</td>
</tr>
<tr>
<td><a href="http://ta.sandart.net">http://ta.sandart.net</a> Meets all standards Full TEI-Download Named Entities, using Identifiers PURLS REST-API, GND-Beacon, LOD/RDF Well-designed user interface with high usability</td>
</tr>
<tr>
<td>No</td>
</tr>
</tbody>
</table>
Appendix B

Chronology: St Augustine and MS XXVIII(26)

The following chronological table is based on (Young et al., 2004, pp. 22-25), (Chadwick, 2004, pp. 328-341), (Hamilton, 2004), (Graham, 2011), (Lawless, 1985), (Weiskotten, 1919) and on the chronological table at the Sant’Agostino: Augustinus Hipponensis website.\(^1\)

<table>
<thead>
<tr>
<th>Date</th>
<th>Augustine</th>
<th>Event</th>
<th>De civitate Dei</th>
<th>MS XXVIII(26)</th>
</tr>
</thead>
<tbody>
<tr>
<td>354</td>
<td>Born on November 13 in Thagaste (Roman province of Numidia; now Souk-Ahras) of a pagan father and a Christian mother.</td>
<td>Adeodatus, son of Augustine, is born.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>372</td>
<td></td>
<td>Teaches rhetoric in Carthage for nine years.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>383</td>
<td></td>
<td>Goes to Rome to teach rhetoric with mistress and son Adeodatus.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^1\)Available at: http://www.augustinus.it/vita/cronologia.htm (Accessed: 21 April 2014).
<table>
<thead>
<tr>
<th>Year</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>384-6</td>
<td>Professor of oratory in Milan. Meets Ambrose, bishop of Milan.</td>
</tr>
<tr>
<td>387</td>
<td>Converts to Christian faith and is baptised by Ambrose at Easter.</td>
</tr>
<tr>
<td></td>
<td>Adeodatus baptised with his father.</td>
</tr>
<tr>
<td>388</td>
<td>Returns to Thagaste in late August/early September</td>
</tr>
<tr>
<td></td>
<td>(and founds his first 'monastery'?)</td>
</tr>
<tr>
<td></td>
<td>Monica, Augustine's mother, dies in Ostia on their way back to Thagaste.</td>
</tr>
<tr>
<td>391</td>
<td>Visits Hippo Regius where he is made presbyter and founds his second monastery.</td>
</tr>
<tr>
<td></td>
<td>Bishop of Hippo is Valerius (from Southern Italy).</td>
</tr>
<tr>
<td>397</td>
<td>Augustine starts writing his Confessiones.</td>
</tr>
<tr>
<td>400</td>
<td>Ordained bishop of Hippo.</td>
</tr>
<tr>
<td>401</td>
<td>Augustine finishes writing his Confessiones.</td>
</tr>
<tr>
<td>410</td>
<td>Fall of Rome.</td>
</tr>
<tr>
<td>413</td>
<td>Augustine writes books 1-5.</td>
</tr>
<tr>
<td>415</td>
<td>Augustine writes 6-10.</td>
</tr>
<tr>
<td>418</td>
<td>Augustine writes books 14-16.</td>
</tr>
<tr>
<td>Year</td>
<td>Event</td>
</tr>
<tr>
<td>------</td>
<td>-------</td>
</tr>
<tr>
<td>426</td>
<td>Augustine writes to Firmus (Epistula ad Firmum); Augustine finishes writing De Doctrina Christiana.</td>
</tr>
<tr>
<td>425-7</td>
<td>Augustine writes books 18-22.</td>
</tr>
<tr>
<td>429</td>
<td>Sack of Hippo by Vandals.</td>
</tr>
<tr>
<td>430</td>
<td>Dies in Hippo.</td>
</tr>
<tr>
<td>431</td>
<td>Siege of Hippo abandoned by the Vandals.</td>
</tr>
<tr>
<td>432-7</td>
<td>Possidius of Calama writes Vita Augustini.</td>
</tr>
</tbody>
</table>
Appendix C

Conservation report of MS

XXVIII(26)

On 23 December 2011, book conservator Dr Alberto Campagnolo was contacted by the author to visit the Biblioteca Capitolare and carry out a conservation assessment of the manuscript.¹ The report, in the Italian original, reads (English translation below):


¹Dr Alberto Campagnolo was a former (2014) book conservator at the Biblioteca Apostolica Vaticana (BAV). At the time of contact, Dr Campagnolo was a full-time PhD student at the University of the Arts with previous work experience in book conservation.
in minuscola umanistica con inchiostro ferro-gallico di colore bruno scuro. Il testo originario del V secolo è stato ritoccato in più punti con inchiostro ferro-gallico nerastro particolarmente acido che ha danneggiato il supporto scrittorio fino a forarlo, specie in carte composte da pergamena sottile o di qualità scadente. Il manoscritto è evidentemente stato danneggiato da umidità e conseguenti attacchi microbiologici che hanno alterato il supporto scrittorio e lasciato tracce, per lo più di colore porpora-rossastro, su circa il 70-80% delle carte, a volte fino a coprire e rendere praticamente illeggibile il testo a occhio nudo. Il manoscritto è stato restaurato, tra il 1923 e il 1927, presso il laboratorio di restauro della Biblioteca Vaticana. L’intervento di restauro è risultato particolarmente invasivo. Le lacune sono state restaurate pergamena con pergamena con evidenti segni di gelatinizzazione della pergamena originaria in un gran numero di carte. Spesso i fori originari del supporto, i fori di tarli, e i fori della rigatura, sono stati rattoppati con un sottile e trasparente materiale di natura proteica e gelatina. Lo stesso materiale è stato applicato con gelatina sui fori lasciati nel testo dall’acidità dell’inchiostro nerastro. In molte carte durante il restauro è stato applicato a pennello uno strato di gelatina per fissare e consolidare gli inchiostri; spesso questo ha causato la gelatinizzazione del supporto di scrittura e la formazione di gore. A causa del degrado della pergamena dovuto agli attacchi microbiologici, lo spessore esiguo di alcune carte, e l’applicazione di gelatina durante il processo di restauro, un numero di carte (10-20% circa) risulta trasparente, specie l’area relativa allo specchio di scrittura, e conseguentemente il testo è di difficile lettura a occhio nudo. Il manoscritto è stato ricucito su tre fettucce, e reincassato nella precedente legatura in piena pergamena rigida (probabilmente del XVIII secolo).

Data la fragilità del supporto scrittorio di molte carte e il cattivo stato generale di conservazione, il manoscritto deve essere considerato a ris-
chio di danni meccanici durante la consultazione. Il manoscritto è stato digitalizzato nel 2003, ma data la qualità della fotografia, la risorsa digitale risulta inadeguata all’uso da parte di studiosi e accademici. Considerata la notevole importanza del testo contenuto nel manoscritto, un nuovo progetto di digitalizzazione, realizzato secondo i più alti standard attuali di conservazione e digitalizzazione, sarebbe auspicabile. La nuova risorsa potrebbe recuperare il testo al momento illeggibile a occhio nudo, cosa che il vecchio progetto del 2003 non ha considerato, e risultare perciò particolarmente utile agli studiosi. Una buona risorsa elettronica potrebbe perciò prevenire danni al manoscritto durante la consultazione e arricchirebbe notevolmente le possibilità di studio al momento criticamente ristrette dovute al cattivo stato di conservazione del manoscritto.

Verona, 23 Dicembre 2011
Alberto Campagnolo

The following is the author’s English translation of the above report:

Parchment codex dated back to the fifth century, containing 254 pages in overall bad condition. The introductory and closing booklets were evidently added at a later date. The writing support varies in quality, from excellent to very low, with several holes and the use of low quality skin parts; the thickness of the parchment also varies from optimal to extremely thin. The original fifth century uncial text is written in dark brown ink, probably gall-ink, which does not appear to have damaged the writing support. The incipit and colophon of the various chapters are written in dark brown and red ink. The introductory and closing gatherings are written in humanistic minuscule in dark brown gall-ink. Parts of the original fifth century text were retraced in a particularly acidic black gall-ink which has damaged the writing support to the point of corrosion, especially
where the parchment is thinnest or of low quality. The manuscript was evidently damaged by humidity and consequent microbiological attacks which affected the writing support and left a purple-red trace on ca. 70-80% of the pages, sometimes even covering the text and making it almost illegible to the human eye. The manuscript was restored between 1923 and 1927 at the conservation laboratory of the Vatican Library. The conservation intervention was rather invasive. The lacunae were restored with parchment. Evident signs of gelatinisation of the original parchment appear in a large number of folia. In several places the original holes in the writing support, produced by worms, as well as the prickings were covered with a thin, transparent, protein-based substance and gelatine. The same substance was applied with gelatine on the holes in the text produced by the acidity of the black ink. During conservation, a layer of gelatine was spread with a brush on many pages in order to bind and consolidate the inks; this has resulted both in the gelatinisation of the writing support and in stains. Owing to the deterioration of the parchment following microbiological attacks and the application of gelatine during conservation, a number of thin pages (10-20%) are almost transparent, especially the written surface, consequently burdening the reading experience. The manuscript was sewn together with three straps and rebound in the previous binding parchment cover (probably 18th century).

Due to the fragility of many pages and the overall bad conservation status, the manuscript must be considered at risk of mechanical damage during consultation. The manuscript was digitised in 2003 but the quality of the photography is not adequate for academic use. Given the noteworthy importance of the text, a new digitisation project, run in accordance with today’s highest conservation and imaging standards, is desirable. The new resource would help recover
the illegible text, something that the 2003 effort did not consider, and would therefore prove useful to scholars. A good digital resource could prevent additional consultation damage and would significantly increase research opportunities currently restricted due to the bad conservation condition of the manuscript.

Verona, 23 December 2011
Alberto Campagnolo