

## Digital Humanities and Qur'ānic Manuscript Studies: New Perspectives and Challenges for Collaborative Spaces and Plural Views

Alba Fedeli

Research Associate, University of Hamburg, Germany

alba.fedeli@uni-hamburg.de

**Abstract:** The use of digital tools has revolutionized academic research in the humanities, changing its approaches, methodologies and outcomes. The phenomenon is comparable to the radical change induced by the introduction of printing and has created a new interdisciplinary area called digital humanities or humanities computing. The application of humanities computing to the study of early Qur'ānic manuscripts can affect the access to and comprehension of these objects in three main aspects. Digital tools change, first, our access to illegible objects through digital images and, second, our editing of manuscript texts by means of tagging systems that produce a structured and searchable/processable archive of the information that manuscripts contain. Last, information technology borrowed from DNA-sequencing tools in biology can be applied to manuscript studies for the encoding of multiple manuscripts and facilitates our understanding by suggesting possible connections between manuscripts. The paper will discuss the methodological innovations – in terms of imaging processing, tagging of text and phylogenetics – of research projects that have already been carried out and the feasibility of establishing a collaborative space for annotating Qur'ānic manuscripts, taking into consideration actors, audiences and purposes.

**Keywords:** Digital humanities; Early Qur'ānic manuscripts; Manuscript CUL Or. 1287; Birmingham Qur'ān; Radiocarbon dating; Phylogenetics; Editing of manuscript texts; Tagging systems; Annotating Qur'ānic manuscripts; Markup language

### العلوم الإنسانية الرقمية والمخطوطات القرآنية: آفاق وتحديات جديدة نحو فضاء للتعاون والتعددية

ألبا فيديلي

باحثة في جامعة هامبورغ - ألمانيا

**ملخص البحث:** إن استخدام الأجهزة الرقمية أحدث ثورة في البحوث الأكاديمية في مجال العلوم الإنسانية، وغير مناهجها ونتائجها. وهذه الظاهرة تعتبر ثورية مقارنة بالتغيير الجذري الذي أحدثته اكتشاف الطباعة، فقد أوجدت الأجهزة الرقمية مجالاً جديداً متعدد التخصصات، يسمى العلوم الإنسانية الرقمية (Digital Humanities) أو حوسبة العلوم الإنسانية (Humanities Computing).

إن تطبيق هذه التقنية في دراسة المخطوطات القرآنية المبكرة يمكن أن يساعد على الوصول إلى هذه المخطوطات وفهمها من ثلاثة جوانب. أولاً: تُغير الأجهزة الرقمية قدرتنا على الوصول إلى المخطوطات غير المقروءة من خلال الصور الرقمية. ثانياً: تُغير تحرير نصوص المخطوطات عن طريق نظام لوضع علامات (Markup)، وبالتالي تنتج قواعد البيانات من المخطوطات. قواعد البيانات هذه هي منظّمة وقابلة للبحث وفي شكل قابل للمعالجة باستخدام الحاسوب. ثالثاً: يتم تطبيق تكنولوجيا المعلومات المستعارة من علم الأحياء في فهم تسلسل الحمض النووي لدراسات المخطوطات من خلال ترميز البيانات في العديد من المخطوطات ويسهل فهمنا باقتراح الروابط المحتملة بين المخطوطات لكي يتسنى لنا معرفة العلاقة المشتركة بينها.

سيناقش بحثي الآثار المنهجية للمشروعات التي تم تحقيقها بالفعل – من حيث معالجة الصور، ووضع علامات على النص، وفهم العلاقة المشتركة بين نصوص المخطوطات عبر استخدام برمجيات شجرة التطور (Phylogenetic software) – ومن حيث جدوى توفر مساحة تعاونية للتعليق على المخطوطات القرآنية، مع الأخذ بعين الاعتبار الجهات الفاعلة والجمهور والأغراض.

**الكلمات المفتاحية:** العلوم الإنسانية الرقمية، المخطوطات القرآنية المبكرة، مخطوطة مكتبة جامعة كامبريدج (المسجلة تحت رقم "MS CUL Or. 1287")، مخطوطة القرآن بجامعة برمنغهام، التأريخ باستخدام الكربون المشع، علم تطور السلالات، تحرير نصوص المخطوطات، التعليق على المخطوطات، لغة ترميزية (نظام لوضع علامات)

Cite this article as: Fedeli A., "Digital Humanities and Qur'ānic Manuscript Studies: New Perspectives and Challenges for Collaborative Spaces and Plural Views", *Journal of College of Sharia and Islamic Studies*, Volume 38, Issue 1, 2020

<https://doi.org/10.29117/jcsis.2020.0260>

© 2020 Fedeli, licensee QU Press. This is an open access article distributed under the terms of the Creative Commons Attribution license CC BY 4.0, which permits unrestricted use, distribution and reproduction in any medium, provided the original work is properly cited.

The use of digital tools pervades every field of academic research in the humanities, changing its approaches, methodologies and outcomes. The phenomenon is comparable to the radical change induced by the introduction of printing and has created a new interdisciplinary area called digital humanities or humanities computing. In the former term, the emphasis is on the digital dimension of tools and the digital display of contents, while the latter refers explicitly to the application of computing, highlighting the structured and thus processable data that scholars use and produce to explore their subjects.

## 1. Introduction

In a period of new interest in Qur'anic manuscripts as physical artifacts and in the readings of these manuscripts, the application of digital technologies to the study of early Qur'anic manuscripts can affect and facilitate access to and comprehension of these objects, which are the oldest extant written evidence of the Qur'anic text. Therefore, manuscript studies is of the utmost importance.

These preserved artifacts are only fragmentary pieces of a whole formed by their total logical sum. Early manuscripts fragments were mainly stored – as in the case of the false ceiling of the Great Mosque of Şan 'ā' – in deposits of unused artifacts rather than being kept in their original place of production and use<sup>(1)</sup>. Because of the particular method of storage in special deposits, early fragments do not necessarily explicitly reveal their origin and thus might be out of context. They very rarely include paratextual elements such as colophons or ownership marks and do not contain explicit key elements regarding their dating and the place where they were elaborated.

In the absence of explicit and absolute references to their dating, early fragments can be dated by means of comparison of their paleographical characteristics with those of contemporary dated materials such as Arabic papyri. In 1958, Adolf Grohmann proposed such a relative chronology placed in relation to internally dated papyri from the first century of the Islamic era<sup>(2)</sup>. A few years earlier, in 1946, Willard Libby developed a new method for dating organic materials by measuring their content of carbon-14, a radioactive isotope of carbon. The new method was based on the interaction in the atmosphere of cosmic ray-derived neutrons with the stable isotope  $^{14}\text{N}$  to form the radionuclide  $^{14}\text{C}$ . Thus, radiocarbon analyses of manuscripts written on parchment calculate the time elapsed between the death of the animal whose skin was used as a writing material and the present time. In fact, the  $^{14}\text{C}$  concentration of living organisms is presumed to be reduced by radioactive decay after their death, while during their life, they have the same  $^{14}\text{C}$  concentration as the atmosphere<sup>(3)</sup>. Despite the fact that radiocarbon analyses can only offer a hypothesis about the date of the animal's death and not the date of the ink's production or the script's execution, their results contribute – with paleographical, art historical and textual analysis – to placing undated Qur'anic fragments within their history and context<sup>(4)</sup>.

(1) They come mainly – as far as we know – from four deposits intended for storing dismissed Qur'anic artifacts, which remained untouched for centuries. The practice of *genizah*-like deposits is described in Joseph Sadan, "Genizah and Genizah-like Practices in Islamic and Jewish Traditions", *Bibliotheca Orientalis*, 1986, 43: 36–58. A contemporary case of a *genizah*-like deposit in Pakistan of contemporary printed copies of the Qur'anic text appeared in a BBC documentary; see BBC, "The Sacred Site for old Korans" (<http://www.bbc.com/news/av/world-asia-43201622/the-tunnels-that-house-old-korans> retrieved on February 11, 2020).

(2) Adolf Grohmann, "The Problem of Dating Early Qur'āns", *Der Islam*, 1958, 33: 213-231.

(3) P. Quentin Dresser, "Radiocarbon dating" in John A. Matthews, ed., *Encyclopedia of Environmental Change* (Los Angeles: SAGE Publications, 2014): 901. Ltd. doi: 10.4135/9781446247501.n3188.

(4) For a complete overview of manuscripts dated by means of  $^{14}\text{C}$  analysis, see Michael Josef Marx and Tobias J. Jocham, "Zu den Datierungen von Koranhandschriften durch die  $^{14}\text{C}$ -Methode", *Frankfurter Zeitschrift für Islamisch-Theologische Studien*, (Koranforschung. Verortung und Hermeneutik), 2015, 2: 9-43 and Marx, Michael Josef and Tobias J. Jocham, "Radiocarbon ( $^{14}\text{C}$ ) Dating of Qur'ān Manuscripts" in Andreas Kaplony – Michael Marx eds., *Qur'ān Quotations Preserved on Papyrus Documents, 7th-10th Centuries and the Problem of Carbon Dating Early Qur'āns* (Leiden: Brill, 2019): 188-221.

Analyses of some manuscripts have brought the documents enormous fame, as in the case of the so-called Birmingham Qur'ān from the Mingana collection<sup>(1)</sup>. The death of the animal used to produce the parchment writing material was placed in the period between 568 and 645 CE with 95.4% probability by a laboratory at the University of Oxford<sup>(2)</sup>, thus confirming that it constitutes part of the corpus of early Qur'ānic manuscripts written in the *hijāzī* script style. Although it is impossible to attribute a precise year/period to the production of the parchment and the execution of the text, all of these fragments form the first extant attestation we have of the written transmission of the text.

Qur'ānic manuscript studies, with the help of digital tools, can lead to the establishment of a well-structured and comprehensive dataset of these artifacts' characteristics, such as linguistic, textual, paleographical and art-historical elements, providing some clues about their production. Moreover, the digital dimension facilitates and fosters collaborative work, which is essential for reaching a new understanding of early manuscripts that is as complete as possible.

Digital tools can change, first, our access to difficult and illegible objects, facilitating our reading through digital images, and second, our editing of manuscript texts by means of tagging systems, thus producing a structured and searchable/processable archive of the information that manuscripts contain. The digital encoding of all of the richness of the manuscripts and their texts allows for the application of programs used in biology to analyze DNA sequences and observe connections between different species. In fact, there are similar patterns in DNA sequences and word sequences, as texts can be reduced to a series of letters as long as manuscripts have readings A or B or C and so on, as DNA sequences do. Moreover, the dynamic and nonlinear dimension of the hypertext may offer an open environment capable of accepting several hypotheses on readings resulting from the interpretation of several scholars in collaborative projects. In Qur'ānic manuscript studies, some objectives following the new digital paradigm have been achieved, while other possibilities – such as an open collaborative space to host a plurality of readings of the manuscript texts – are something we are only envisaging based on models of projects carried out in other textual traditions.

## **2. Digital editions as multilayered open texts and processable/processed data: tagging the plurality of readings**

Regarding the use of digital tools to access the physicality of manuscripts and their text, digital images are indispensable in recovering damaged or illegible objects. Digital images taken using the application of ultraviolet light and their processing by means of software has made possible the unveiling of – for example – the underwriting of the Cambridge Qur'ānic palimpsest that was cancelled to obtain recycled parchment material for assembling a new codex<sup>(3)</sup>.

In the digital paradigm, the intrinsic nature of its dimensions frees manuscripts' edition from the constraints

- 
- (1) The importance of the Birmingham Qur'ān, with its analysis and the rediscovery of its story, is described in Alba Fedeli, "Early Qur'ānic Manuscripts, their Text, and the Alphonse Mingana Papers Held in the Department of Special Collections of the University of Birmingham" (Unpublished PhD thesis, School of Philosophy, Theology and Religion, University of Birmingham, UK, 2015).
  - (2) Sample OxA-29418 parchment, ?goat, MS1572 Cadbury Research Library,  $\delta^{13}\text{C} = -21.0\text{‰}$   $1456 \pm 21$ , in T. F. G. Higham, C. Bronk Ramsey, D. Chivall, J. Graystone, D. Baker, E. Henderson and P. Ditchfield, "Radiocarbon Dates from the Oxford AMS System: Archaeometry Datelist 36", *Archaeometry*, 2018, 60 (3): 628-640, 634.
  - (3) The images of the Cambridge (Mingana-Lewis) palimpsest – available online in the Cambridge Digital Library <https://cudl.lib.cam.ac.uk/collections/minganalewis/1> – are just an expedient to display an interpretation of the palimpsest text. Given the perspective and methodology applied in processing the Cambridge palimpsest images, the resulting "reconstruction" is a hypothetical retracement of the underwriting of the manuscript or, in other words, an interpretation of its text.

of the linear, circumscribed and static page of the printed text<sup>(1)</sup>. Thus, the digital edition can host and express a mobile, multilayered text, although the condition for production of a digital edition is that the manuscript text be processable rather than merely appropriate for digital display. The multilayered nature of the text is a consequence of the possibilities given by the environment, while the existence of processable data is the necessary precondition for the elaboration of a digital edition. Computational editing allows scholars to produce a structured interpretation of manuscript data. As their reading is structured, it is processable<sup>(2)</sup>. Moreover, digital editing implies the possibility of editing the multiple layers of a manuscript and its variances. As early as the end of the 1980s, before the release of pioneering digital editions, Cerquiglini defended movement and fluidity in manuscript culture. According to the scholar, the electronic form is momentary and basically modifiable, like the medieval text, thus making the electronic medium the solution to the problem of rendering manuscript movement<sup>(3)</sup>.

The digital environment has the capacity to admit editions that encompass the richness of all the available manuscript witnesses of a text, including all the details of each witness and its highly layered structure. The theory of digital scholarly editing can capture the variance<sup>(4)</sup>, presenting and reproducing the variant states of the manuscript text in their entirety and all of the manuscript evidence without confining this richness to a decontextualized *apparatus criticus* and without normalizing the linguistic situation expressed by scribes in the written textual tradition of the Qur'ān. Thus, for example, all the available and known Qur'anic manuscripts can be presented in full, acknowledging the specific value of every reading. Moreover, the scholar encoders can transcribe and include their editorial notes and all the details of each single manuscript, which cannot be reproduced in a printed edition. The introduction of printing eliminated elements connected with the materiality of the manuscript object. As observed by Ginzburg, the introduction first of writing and then of printing dematerialized the text. Therefore, textual scholarship – in a printing environment – can only take into consideration textual features that can be reproduced<sup>(5)</sup>.

Two – among many other – situations encountered in early Qur'anic manuscripts can representatively illustrate the dematerialization of their content in printed editions. First, manuscripts have traces of production, use and additions through time and through several stages in their process. Thus, first hands and later hands as well as corrections, ameliorations and additions are a common situation in early Qur'anic manuscripts, whereas their printed editions tend to present one stage of the writing process, i.e., the possible original stage or the "correct(ed)" stage. The second example of dematerialization of the text is the loss of information carried by the coexistent double presence of diacritical marks in a manuscript dimension, where its possible parallel printed edition renders just one of the two pieces of information or presents one reading as the main one and confines the other to a footnote. For example, in a manuscript artifact, both *li-yuṭahhira-kum*, i.e., "so that He may purify you", and *li-nuṭahhira-kum*, i.e., "so that We may purify you", can be expressed in a single word/

(1) Buzzetti highlights the possibility that a strongly embedded markup inherits many of a text's properties, including the character sequence and its linear order. See Dino Buzzetti, "Digital Representation and the Text Model", *New Literary History*, 2002, 33 (1): 61-87, 67.

(2) See, for example, Dino Buzzetti, "Digital Editions and Text Processing", in Kathryn Sutherland and Marilyn Deegan eds., *Text Editing, Print and the Digital World* (Abingdon: Routledge, 2016) 2<sup>nd</sup> ed. (Farnham: Ashgate Publishing, 2009): 45-61.

(3) Bernard Cerquiglini, *Éloge de la variante: histoire critique de la philologie* (Paris: Les Éditions du Seuil, 1989): 105-116.

(4) This is the approach of genetic criticism. This French movement of the 1960s replaced the concept of the critical edition with the idea of a "genetic dossier" that "includes all surviving documents of the writing process, each retaining its own integrity". See Neil Fraistat and Julia Flanders, "Introduction" in Neil Fraistat and Julia Flanders eds., *The Cambridge Companion to Textual Scholarship* (Cambridge: Cambridge University Press, 2013): 1-15, 6.

(5) Carlo Ginzburg, "Spie, Radici di un paradigma scientifico" in Aldo Gargani ed., *Crisi della ragione* (Torino: Einaudi, 1979): 72. The dematerialization of the text is mentioned in Giuseppe Bompreszi and Vincenzo Fano, "Considerazioni epistemologiche sulle scienze storico-filologiche", *Studi Urbinate, B - Scienze umane e sociali*, 2011, 81: 21-44.

entity with the same identical value insofar as both diacritical signs below and above the consonantal skeleton are traced by the same hand with the same ink<sup>(1)</sup>. In a digital environment, scholars can tag such a situation with a markup system that encodes the presence of two coexisting and equal readings in a single entity.

However, it has to be noted that – even in a multistructured and multilayered digital edition of a manuscript – the use of the Unicode font forces scholars to adapt their encoding to the limited possibilities reproduced by the Unicode set of characters, which have merely inherited the system of typography born in a non-Arabic environment. Considering the abovementioned example, the letter *nūn* in *li-nuṭahhira-kum* can be represented in early manuscripts by an unspecified – or ambiguous – denticle not possessing a diacritical mark above it or by a denticle specified by its diacritical mark above it. In the transcription of the word, scholars lose the specific characteristics of the Arabic writing system and accept the compromise of reproducing what the Unicode font offers, i.e., a dotless letter *bā'* expressed by the code 066E. Nevertheless, the dotless letter *bā'* 066E is as ambiguous as the letter *nūn* with its code 0646 is in transcribing the dotless denticle of the manuscript. The former displays a dotless denticle carrying the wrong information, while the second displays and carries an incorrect situation. The transcribed text is dematerialized from its physicality as expressed in manuscript culture. Cases like this pose problems when – for example – scholars want to compare transcribed words with the help of programs. Such problems could be solved by splitting the dematerialized typographical idea of a unique character carrying a single entity of denticle and diacritic into two distinct entities placed horizontally one after the other, i.e., denticle and diacritic. The materiality of the Arabic writing system conveys two pieces of information rather than a single piece made of two assembled elements<sup>(2)</sup>.

Although accepting – initially inevitable – compromises inherited from the printing paradigm, scholarly digital editions do not have the limitations of the linear and static transcription of the manuscript text and features. Scholars can tag the text and add elements after the analysis by using a markup language to encode and share their edition, for example, SGML, XML, HTML or JSON. The markup and tagging of the text are not only instruments for achieving semiotic tasks but also represent and produce a separate semiotic dimension<sup>(3)</sup>.

Thus, for example, for the digital edition of the Cambridge Qur'ānic palimpsest, available in the Cambridge Digital Library, I tagged the manuscript using opening and closing elements, entities and markers of positioning in the work and in the physical object based on TEI P5 converted to XSL and XML<sup>(4)</sup>. In the particular case of the difficult palimpsest object, the digital edition provided the tools for documenting the interpretation of the images and their text. This produces a digital documentary edition, a born-digital edition. Such a digital documentary edition tags the manuscript, its text, and scholars' observations. Editorial notes and tags encode the manuscript and its text in categories that build and produce a database, i.e., a searchable/processable archive of information that the manuscript contains<sup>(5)</sup>.

(1) In Q.8:11, a manuscript of the National Library of Russia reads both – and simultaneously – *li-yuṭahhira-kum* ("so that He may purify you") and *li-nuṭahhira-kum* ("so that We may purify you"), with both diacritical markers traced by the first hand in the same ink color. Both readings have the same value. This argument has been illustrated in Alba Fedeli and Andrew Edmondson, "Early Qur'anic Manuscripts and their Networks: a Phylogenetic Analysis project", a paper presented at the "Paleo-Qur'anic Manuscripts: State of the Field" Conference, Budapest, May 2017.

See the printed form (*li-yuṭahhira-kum*) of the manuscript text in the edition by Éléonore Cellard and Sabrina Cimiotti, *Codex Amrensis I*. Documenta Coranica, Manuscripta, vol. 1 (Leiden: Brill, 2018): 316-317.

(2) Thomas Milo is working in this direction in developing his "DecoType project".

(3) Domenico Fiorimonte, *Scrittura e filologia nell'era digitale* (Torino: Bollati Boringhieri, 2003): 219.

(4) See for example XML encoding for manuscript transcription in Hugh A.G. Houghton, "The Electronic Scriptorium: Markup for New Testament Manuscripts" in Claire Clivaz, David Hamidovic and Andrew Gregory, eds., *Digital Humanities in Biblical, Early Jewish and Early Christian Studies* (Leiden: Brill, 2014): 31-60, 38 and foll.

(5) The very concise and precise definition of a digital edition as "an image of the text" and "processable data" is given in Buzzetti, "Digital Editions and Text Processing", 45-46.

Thus, I encoded the text of the Cambridge palimpsest introducing editorial notes referring to the writing process or the use of the text itself, for example "codicology structure", "end of verse", "script", "writing process", "layout", "parchment", and "use of the text". Moreover, the text was structured through the insertion of editorial notes describing particular situations of the text and its variant readings, focusing on their linguistic implications. For example, editorial notes refer to situations that involve – among others – orthography, phonetics, morphology, syntax and lexicon. In the online edition, a transcribed word is accompanied by a pop-up window showing the user that the word is connected – for example – with the phenomenon of the absence of cases, as far as we know, from early Arabic grammar. The user can read about the "absence of cases", and while the category "morphology" is invisible to the user, this category has been used to build a database of information, in this case from a linguistic point of view. Scholars can structure their database according to their research questions. Focusing on the linguistic connotations of the specific manuscript readings allows for the exploration of the interference of the linguistic system of scribes in their work of copying/writing down a text, thus revealing possible significant traces of the historical context of these early artifacts<sup>(1)</sup>.

### 3. Computing and displaying the digital data: Processing the plurality of manuscript readings and accepting the plurality of editorial views

Differences in writing manuscripts are an intrinsic factor resulting from the act of copying or writing down a text. Scholars can analyze several manuscripts to compare their texts and paratextual features (i.e., agreements and disagreements) to understand the relationships between them. This kind of analysis is of great importance, as it can provide clues for grouping manuscripts and tracing a new picture of fragments about whose production we have no details.

The analysis of these possible connections is an unexplored field in Qur'anic studies after works based on a traditional stemmatic approach by Theodor Nöldeke in 1860, Michael Cook in 2004 and Behnam Sadeghi in 2010<sup>(2)</sup>. These previous attempts to understand the family trees of the witnesses were based on readings that Islamic sources attribute to early manuscripts, no longer extant<sup>(3)</sup>, while the exploration of the manuscript tradition is a new research direction in Qur'anic manuscript studies. In the last few years, the use of computer programs introduced the possibility of a "new stemmatic" based on a cladistics method in investigating the relationship between manuscripts<sup>(4)</sup>. The interdisciplinary approach given by philological analysis and the use of phylogenetic software developed in biology offers an innovative instrument for studying manuscript traditions. In fact, the use of computer programs for estimating mutations in DNA and their propagation (phylogenetics or phylogenetic systematics) has assisted textual scholars in handling relationships among manuscripts (computer-aided stemmatic or phylomemetics)<sup>(5)</sup>. The methodology has already been tested on a small portion of the Qur'anic text, focusing on the establishment of a system for encoding the text in a manner suitable for phylogenetic analysis software. After several attempts that were continuously implemented and changed, I encoded the manuscript data in an Excel spreadsheet, later converted into a NEXUS file, suitable for analysis through phylogenetic software<sup>(6)</sup>. I structured the manuscript data by distinguishing among the three

(1) See Segre, *Semiotica filologica. Testo e modelli culturali* (Torino: Einaudi, 1979) and Bomprezzi and Fano, "Considerazioni epistemologiche sulle scienze storico-filologiche", 38.

(2) Further analysis in Fedeli and Edmondson, "Early Qur'anic Manuscripts and their Networks: A Phylogenetic Analysis project".

(3) Benham Sadeghi and Uwe Bergmann, "The Codex of a Companion of the Prophet and the Qur'ān of the Prophet", *Arabica*, 2010, 57: 343-436, included a few leaves of the Sanaa palimpsest in their analysis.

(4) Peter M.W. Robinson and Robert J. O'Hara, "Report on the Textual Criticism Challenge 1991," *Bryn Mawr Classical Review* 1992, 3: 331-337.

(5) C. J. Howe and H. F. Windram, "Phylomemetics—Evolutionary Analysis beyond the Gene", *PLoS Biology*, 9 (5), 2011, e1001069. doi:10.1371/journal.pbio.1001069, retrieved on February 10, 2020.

(6) Andrew Charles Edmondson, *Phylogenetic Analysis of the Gospel of John Using BlueBEAR* (CCB Seminar 2016), [https://www.academia.edu/28678521/Phylogenetic\\_analysis\\_of\\_the\\_Gospel\\_of\\_John\\_using\\_BlueBEAR\\_CCB\\_Seminar](https://www.academia.edu/28678521/Phylogenetic_analysis_of_the_Gospel_of_John_using_BlueBEAR_CCB_Seminar), retrieved on February 10, 2020. See also Andrew Charles Edmondson, "An Analysis of the Coherence-Based Genealogical Method using Phylogenetics" (Unpublished PhD thesis, School of Philosophy, Theology and Religion, University of Birmingham, UK, 2019).

main elements inserted in the EXCEL file, namely, Qur'ānic text, paratextual elements and ends of verses. Each point of disagreement in the manuscripts is characterized by the insertion of abbreviations that indicate the reading typology, according mostly to the encoding system described above for the Cambridge palimpsest. Such a structure for transcribing the data allows filters to be applied when we run the software. In fact, for example, an orthographic variant could be genetically irrelevant, whereas a morphological variant could be genetically relevant and indicate a close connection between two manuscripts. The encoded classification of the variant typologies assures the possibility of applying and experimenting with filters, thus excluding or including certain typologies of variants for tracing possible connections and networks.

Drawing this distinction between variants to weight their typologies is an essential part of the stemmatic analysis of manuscript texts, as scholars can take a weighted or a nonweighted phylogenetic approach in applying computational stemmatic analysis. In a recent study, T.L. Andrews and C. Macé argue that even the most insignificant changes have "some text-genealogical significance that should not be discounted". Moreover, once scholars have proposed hypotheses about a possible group of manuscripts that are genetically closely related, the high frequency of "mistakes" (i.e., genetically irrelevant variants) in that group can add credibility to the stemmatic hypothesis<sup>(1)</sup>.

The results from the testing phase in using phylogenetic software allowed us to envisage the planning of a larger – and possibly collaborative – project to draw a more detailed story of the early fragments understood as a corpus rather than as isolated items. Computer programs calculate and suggest connections between manuscripts based on their agreements and disagreements so that scholars can evaluate and verify the more plausible connections and groupings. However, the analysis could be implemented with a more important role played by the computer in attributing the typologies of variants through natural language computing technology for the automatic analysis of texts and possibly adding semantic analysis<sup>(2)</sup>. Moreover, the computer could easily solve the problem of the ambiguity of the text due to the peculiar system of early Arabic writing as described above. The development of a different type of system that splits a "single character" into two or three different pieces of information would first reflect the physically rich system of Arabic writing and second allow for automatic inference and analysis by the computer when comparing texts from several manuscripts.

Digital technologies applied to the humanities have shaped a new approach to studying manuscripts and texts guided by a digital paradigm in its theory, method and practice<sup>(3)</sup>. This paradigm allows the history of the transmission to be represented, including all the witnesses and the entirety of each single witness with its (so-called) mistakes that are relevant/significant for analyzing the tradition and social dimensions of the text and do not necessarily have to be regularized<sup>(4)</sup>. The epistemological paradigm of digital philology does not rest upon a generalization of the manuscript process but accounts for the singular specific cases. They – the specific cases – can be accepted in a digital environment.

More significantly, the hypertext structure admits the recognition of a multiplicity of views and interpretations,

- 
- (1) Philipp A. Maas, "Computer Aided Stemmatics — The Case of Fifty-Two Text Versions of Carakasamhitā Vimānasthāna 8.67-157", *Wiener Zeitschrift für die Kunde Südasiens / Vienna Journal of South Asian Studies*, (Text Genealogy, Textual Criticism and Editorial Technique), 2009-2010, 52/53: 63-119, 66, states, "Mistakes that can be easily corrected do not reveal the genealogical relationship of manuscripts in their own right. If, however, these variants occur frequently within a genealogically closely related group of witnesses, they add credibility to the stemmatic hypothesis".
  - (2) See, for example, the *Qurānic Arabic Corpus* project. This is part of a PhD research project into Arabic language computing at the University of Leeds, focusing on the application of Arabic computing language technology to the Qur'ān to achieve morphological and syntactic analysis.
  - (3) The digital paradigm is described in Patrick Sahle, "What is a Scholarly Digital Edition?" in Matthew James Driscoll and Elena Pierazzo eds., *Digital Scholarly Editing. Theories and Practices* (Cambridge: Open Book Publishers, 2016): 19-39.
  - (4) About the loss of characteristics of the graphic and phonetic system and regularization, see Francesca Tomasi, *Metodologie informatiche e discipline umanistiche* (Roma: Carocci editore, 2008): 162 and 163.

taking a pluralist attitude. The recognition of the plurality of views and the subsequent collaborative implications would be extremely effective in enriching our knowledge on Qur'anic manuscripts. The work of reading manuscripts is open to a plurality of interpretations from the point of view of the receptors. The results of the philological work are hypotheses from an epistemological point of view<sup>(1)</sup>, and different scholars can interpret Qur'anic manuscripts in different ways. Moreover, these artifacts give room to pluralistic positions from the point of view of their producers as well because of additions at later stages of the manuscript's life and, in particular, the mechanism of simultaneously placing two positions for diacritics and vowel dots.

Thus, based on the model of the *Workspace for Collaborative Editing* project carried out to support the production of the *Editio Critica Maior* of the Greek New Testament by editors based all over the world<sup>(2)</sup>, we envisage the planning of an online editor for Qur'anic manuscripts considering a hypertext structure that could accept and display different interpretations by different scholars. The *Workspace* consists of a tool whose interface replaces the markup system with dialogue boxes that hide the complex tagging system<sup>(3)</sup> and reduce the percentage of mistakes from the editors. Moreover, the transcription editor "enable[s] dispersed collaborators to work with differing operating systems and contribute directly to the central data store"<sup>(4)</sup>. The *Workspace* was specifically planned for Greek New Testament manuscripts and was later customized for an edition of Avestan texts for possible development for use with other textual traditions<sup>(5)</sup>.

A possible alternative for creating a collaborative space is the multiversion document (MVD) model. It represents a text as a set of merged versions in a single digital entity, creating a network of texts and thus allowing a network of editions realized by different scholars' desiderata of collaborative creation, verification and refinement<sup>(6)</sup>.

The digital interface admits a single edition representing "the disparate views of many editors, allowing the reader to experience the very debatedness of the text as part of its presentation" and the "sense of textual plurality"<sup>(7)</sup>.

The intrinsic nature of digital tools offers all the instruments for crossing possible boundaries between scholars working independently and with different methodological approaches. The conference on *Contemporary*

- 
- (1) August Boeckh, as mentioned in Bompreszi and Fano, "Considerazioni epistemologiche sulle scienze storico-filologiche", 41.
  - (2) Hugh A.G. Houghton and Catherine J. Smith, "Digital Editing and the Greek New Testament" in Claire Clivaz, Paul Dilley and David Hamidović eds., *Ancient Worlds in Digital Culture* (Leiden: Brill, 2016): 110-127 and Houghton, "The Electronic Scriptorium".
  - (3) Online transcription editors used within a web browser eliminate the problem whereby "the verbose character of the markup makes it [XML] very inefficient for transcribers to work directly in this encoding", see Houghton, "The Electronic Scriptorium", 36.
  - (4) Hugh A.G. Houghton, Catherine J. Smith and Martin Sievers, "The Workspace for Collaborative Editing" in *Digital Humanities 2014 Conference Abstracts*, EPFL–UNIL, Lausanne, Switzerland, 8–12 July 2014.
  - (5) Another project conceived of as a collaborative work by means of digital tools and its digital dimension is the project "Philosophy on the Border of Civilizations and Intellectual Endeavours: Towards a Critical Edition of the Metaphysics (*Ilāhiyyāt of Kitāb al-Šifā'*) of Avicenna (Ibn Sīnā)" based at the Scuola Normale Superiore of Pisa and the School of Advanced Studies of Lucca, see <https://www.avicennaproject.eu/#/> retrieved on February 11, 2020.
  - (6) See Desmond Schmidt and Robert Colomb, "A data structure for representing multi-version texts online", *International Journal of Human - Computer Studies*, 2009, 67 (6): 497-514. The MVD model is based on a standoff markup system that allows different types of linguistic analysis, which is another desideratum in Qur'anic manuscript studies. See Desmond Schmidt, "The role of markup in the digital humanities", *Historical social research*, 2012, 37 (3): 125-146. Moreover, this representation has the advantage of facilitating comparison between manuscripts and handling partial versions of a text such as early manuscript fragments (see Schmidt and Colomb, "A data structure", 504 and 508). I thank Dino Buzzetti for suggesting to me the MVD alternative.
  - (7) Fraistat and Flanders, "Introduction", 13.



*Western Trends in Qur'anic Studies. The Problem of Objectivity and Bias: An Epistemological Perspective* has been the ideal location for illustrating the possible epistemological outcomes of a platform that could admit a plurality of readings in manuscripts as written by scribes and a plurality of interpretations in manuscripts as read by editors based all over the world and working with different operating systems in a plurality of transparently displayed data. Qur'anic textual scholarship is no longer a specialized field of hypotheses open only to editors but has significance for a larger public clamoring for a publicly visible, open and accessible tool showing transparently the plurality of views. Two examples are indicative of the implications of such an approach. The first example is the 1914 edition of the Cambridge palimpsest, edited and published by Alphonse Mingana and Agnes Smith Lewis in 1914<sup>(1)</sup>. Immediately, and for a long time afterwards, Mingana was ferociously criticized for the palimpsest edition, which was considered unreliable and incorrect. Leaving aside the discussion about its being correct or incorrect<sup>(2)</sup>, the reader of the 1914 edition would have liked – and would like – to have the possibility to see the manuscript images, the Mingana-Lewis reading and the readings proposed by the critics, all in a unique environment representing the alternative viewpoints. In the digital edition of the Cambridge Digital Library, images sit beside my parallel interpretation. Although it was technically impossible to obtain the display of multiple editions as an outcome of my research project, I clearly included the concept of the plurality of views at the very basis of my encoding. It is important to distinguish and admit actual possibilities and future plans, but the limitations of the display cannot change the awareness of the distinction between encoding and interface. I marked all the points at which I read the manuscript text differently from Mingana and Lewis by inserting editorial notes referring to the "1914 Mingana-Lewis edition". A second example of desiderata for the transparent display of multiple editions is probably the work that emerged about the highly discussed palimpsest of Sanaa<sup>(3)</sup>. In a period of rapid change in the technical possibilities of the digital environment, it would be useful not to present one edition as a final statement but to display all of them as alternative editions, as a "field of intelligently structured possibility"<sup>(4)</sup>. This "structured possibility" would be to the advantage of our comprehension of the Sanaa palimpsest.

The methodological implications of the research projects already realized are challenging – in terms of image processing, text tagging and phylogenetics – and the feasibility of a future collaborative space for annotating

- 
- (1) Alphonse Mingana and Agnes Smith Lewis, *Leaves from Three Ancient Qurāns Possibly pre-'Othmānic with a List of their Variants* (Cambridge: Cambridge University Press, 1914).
  - (2) I explored the polemics over Mingana's edition mainly in Alba Fedeli, "The Digitization Project of the Qur'anic Palimpsest, MS Cambridge University Library Or. 1287, and the Verification of the Mingana-Lewis Edition: Where is salām?", *Journal of Islamic Manuscripts*, 2011, 2 (1): 100-117 and Fedeli, "Early Qur'anic Manuscripts, their Text, and the Alphonse Mingana Papers".
  - (3) Editions of the manuscript text have been published by Benham Sadeghi and Mohsen Goudarzi, "Ṣan'ā' I and the Origins of the Qur'ān", *Der Islam*, 2012, 87: 1-129, Asma Hilali, *The Sanaa Palimpsest. The Transmission of the Qur'an in the First Centuries AH* (Oxford: Oxford University Press, 2017), and several articles by Elisabeth Puin, "Ein früher Koranpalimpsest aus Ṣan'ā' (DAM 01-27.1)" in Markus Groß – Karl-Heinz Ohlig, eds., *Schlaglichter. Die beiden ersten islamischen Jahrhunderte* (Berlin: Verlag Hans Schiler, 2008): 461-493; "Ein früher Koranpalimpsest aus Ṣan'ā' (DAM 01-27.1). Teil II" in Markus Groß – Karl-Heinz Ohlig, eds., *Vom Koran zum Islam: Schriften zur Frühen Islamgeschichte und zum Koran* (Berlin: Verlag Hans Schiler, 2009): 523-581; "Ein früher Koranpalimpsest aus Ṣan'ā' (DAM 01-27.1). Teil III: Ein nicht-'uṭmānischer Koran" in Markus Groß – Karl-Heinz Ohlig, eds., *Die Entstehung einer Weltreligion I. Von der koranischen Bewegung zum Frühislam* (Berlin: Verlag Hans Schiler, 2010): 233-305; "Ein früher Koranpalimpsest aus Ṣan'ā' (DAM 01-27.1). Teil IV: Die scriptio inferior auf den Blättern 17, 18 und 19 der Handschrift DAM 01-27.1 (Sure 9:106-Ende, dann 19:1-67 und weiter)" in Markus Groß – Karl-Heinz Ohlig, eds., *Die Entstehung einer Weltreligion II. Von der koranischen Bewegung zum Frühislam* (Berlin: Verlag Hans Schiler, 2011): 311-402; "Ein früher Koranpalimpsest aus Ṣan'ā' (DAM 01-27.1). Teil V: Die scriptio inferior auf den Blättern 14 und 15 sowie Auseinandersetzung mit den Thesen und der Edition des Koranpalimpsests von Behnam Sadeghi und Mohsen Goudarzi" in Markus Groß – Karl-Heinz Ohlig, eds., *Die Entstehung einer Weltreligion III. Die heilige Stadt Mekka – eine literarische Fiktion* (Berlin: Verlag Hans Schiler, 2014): 477-618. The *Corpus Coranicum* edition is forthcoming.
  - (4) The expression "field of intelligently structured possibility" is in Fraistat and Flanders, "Introduction", 13.

Qur'anic manuscripts and displaying several editions calls for the "right to pluralism" that the conference on *Contemporary Western Trends in Qur'anic Studies. The Problem of Objectivity and Bias: An Epistemological Perspective* held at the College of Sharia and Islamic Studies of Qatar wanted to foster. The new paradigm offers a new epistemological approach in Qur'anic manuscript studies, accepting a great quantity of varied hypotheses and crossing boundaries, offering a transnational and transcultural system<sup>(1)</sup>, beyond discussions about different approaches in a Muslim or Orientalist culture as well as in the East or the West. This is how textual criticism in Qur'anic manuscript studies may develop further in the future.

---

(1) Jerome McGann, "Why Digital Textual Scholarship Matters; Or, Philology in a New Key". Coda in Neil Fraistat and Julia Flanders eds., *The Cambridge Companion to Textual Scholarship*, (Cambridge: Cambridge University Press, 2013): 274-288, 275.

## References

- Bomprezzi, Giuseppe and Vincenzo Fano, "Considerazioni epistemologiche sulle scienze storico-filologiche", *Studi Urbinati, B - Scienze umane e sociali*, 2011, 81: 21-44.
- Buzzetti, Dino, "Digital Editions and Text Processing", in Kathryn Sutherland and Marilyn Deegan eds., *Text Editing, Print and the Digital World* (Abingdon: Routledge, 2016) 2<sup>nd</sup> ed. (Farnham: Ashgate Publishing, 2009): 45-61.
- Buzzetti, Dino, "Digital Representation and the Text Model", *New Literary History*, 2002, 33 (1): 61-87.
- Cellard, Éléonore and Sabrina Cimiotti, *Codex Amrensis I. Documenta Coranica, Manuscripta*, vol. 1 (Leiden: Brill, 2018).
- Cerquiglioni, Bernard, *Éloge de la variante: histoire critique de la philologie* (Paris: Les Éditions du Seuil, 1989).
- Dresser, P. Quentin, "Radiocarbon dating" in John A. Matthews, ed., *Encyclopedia of Environmental Change* (Los Angeles: SAGE Publications, 2014): 901. Ltd. doi: 10.4135/9781446247501.n3188.
- Edmondson, Andrew Charles, *Phylogenetic Analysis of the Gospel of John Using BlueBEAR* (CCB Seminar 2016), [https://www.academia.edu/28678521/Phylogenetic\\_analysis\\_of\\_the\\_Gospel\\_of\\_John\\_using\\_BlueBEAR\\_CCB\\_Seminar](https://www.academia.edu/28678521/Phylogenetic_analysis_of_the_Gospel_of_John_using_BlueBEAR_CCB_Seminar), retrieved on February 10, 2020.
- Edmondson, Andrew Charles, "An Analysis of the Coherence-Based Genealogical Method using Phylogenetics" (Unpublished PhD thesis, School of Philosophy, Theology and Religion, University of Birmingham, UK, 2019).
- Fedeli, Alba, "Early Qur'ānic Manuscripts, their Text, and the Alphonse Mingana Papers Held in the Department of Special Collections of the University of Birmingham" (Unpublished PhD thesis, School of Philosophy, Theology and Religion, University of Birmingham, UK, 2015).
- Fedeli, Alba, "The Digitization Project of the Qur'ānic Palimpsest, MS Cambridge University Library Or. 1287, and the Verification of the Mingana-Lewis Edition: Where is salām?", *Journal of Islamic Manuscripts*, 2011, 2 (1): 100-117.
- Fiorimonte, Domenico, *Scrittura e filologia nell'era digitale* (Torino: Bollati Boringhieri, 2003).
- Fraistat, Neil and Julia Flanders, "Introduction" in Neil Fraistat and Julia Flanders eds., *The Cambridge Companion to Textual Scholarship* (Cambridge: Cambridge University Press, 2013): 1-15.
- Ginzburg, Carlo, "Spie, Radici di un paradigma scientifico" in Aldo Gargani ed., *Crisi della ragione* (Torino: Einaudi, 1979).
- Grohmann, Adolf, "The Problem of Dating Early Qur'āns", *Der Islam*, 1958, 33: 213-231.
- Higham, T. F. G., C. Bronk Ramsey, D. Chivall, J. Graystone, D. Baker, E. Henderson and P. Ditchfield, "Radiocarbon Dates from the Oxford AMS System: Archaeometry Datelist 36", *Archaeometry*, 2018, 60 (3): 628-640.
- Hilali, Asma, *The Sanaa Palimpsest. The Transmission of the Qur'an in the First Centuries AH* (Oxford: Oxford University Press, 2017).
- Houghton, Hugh A.G., "The Electronic Scriptorium: Markup for New Testament Manuscripts" in Claire Clivaz, David Hamidovic and Andrew Gregory, eds., *Digital Humanities in Biblical, Early Jewish and Early Christian Studies* (Leiden: Brill, 2014): 31-60.
- Houghton, Hugh A.G. and Catherine J. Smith, "Digital Editing and the Greek New Testament" in Claire Clivaz, Paul Dilley and David Hamidović eds., *Ancient Worlds in Digital Culture* (Leiden: Brill, 2016): 110-127.
- Houghton, Hugh A.G., Catherine J. Smith and Martin Sievers, "The Workspace for Collaborative Editing" in *Digital Humanities 2014 Conference Abstracts*, EPFL–UNIL, Lausanne, Switzerland, 8–12 July 2014.
- Howe, Christopher J. and Heather F. Windram, "Phylomemetics—Evolutionary Analysis beyond the Gene", *PLoS Biology*, 9 (5), 2011, e1001069. doi:10.1371/journal.pbio.1001069, retrieved on February 10, 2020.
- Maas, Philipp A., "Computer Aided Stemmatology — The Case of Fifty-Two Text Versions of Carakasamhitā Vimānasthāna 8.67-157", *Wiener Zeitschrift für die Kunde Südasien / Vienna Journal of South Asian Studies*, (Text Genealogy, Textual Criticism and Editorial Technique), 2009-2010, 52/53: 63-119.
- Marx, Michael Josef and Tobias J. Jocham, "Zu den Datierungen von Koranhandschriften durch die 14C-Methode", *Frankfurter Zeitschrift für Islamisch-Theologische Studien*, (Koranforschung. Verortung und Hermeneutik), 2015, 2: 9-43.
- Marx, Michael Josef and Tobias J. Jocham, "Radiocarbon (<sup>14</sup>C) Dating of Qur'ān Manuscripts" in Andreas Kaplony – Michael Marx eds., *Qur'ān Quotations Preserved on Papyrus Documents, 7th-10th Centuries and the Problem of Carbon Dating Early Qur'āns* (Leiden: Brill, 2019): 188-221.

- McGann, Jerome, "Why Digital Textual Scholarship Matters; Or, Philology in a New Key". Coda in Neil Fraistat and Julia Flanders eds., *The Cambridge Companion to Textual Scholarship* (Cambridge: Cambridge University Press, 2013): 274-288.
- Mingana, Alphonse and Agnes Smith Lewis, *Leaves from Three Ancient Qur'ans Possibly pre-'Othmânic with a List of their Variants* (Cambridge: Cambridge University Press, 1914).
- Puin, Elisabeth, "Ein früher Koranpalimpsest aus Şan'ā' (DAM 01-27.1)" in Markus Groß – Karl-Heinz Ohlig, eds., *Schlaglichter. Die beiden ersten islamischen Jahrhunderte* (Berlin: Verlag Hans Schiler, 2008): 461-493.
- Puin, Elisabeth, "Ein früher Koranpalimpsest aus Şan'ā' (DAM 01-27.1). Teil II" in Markus Groß – Karl-Heinz Ohlig, eds., *Vom Koran zum Islam: Schriften zur Frühen Islamgeschichte und zum Koran* (Berlin: Verlag Hans Schiler, 2009): 523-581.
- Puin, Elisabeth, "Ein früher Koranpalimpsest aus Şan'ā' (DAM 01-27.1). Teil III: Ein nicht-'uṭmānischer Koran" in Markus Groß – Karl-Heinz Ohlig, eds., *Die Entstehung einer Weltreligion I. Von der koranischen Bewegung zum Frühislam* (Berlin: Verlag Hans Schiler, 2010): 233-305.
- Puin, Elisabeth, "Ein früher Koranpalimpsest aus Şan'ā' (DAM 01-27.1). Teil IV: Die scriptio inferior auf den Blättern 17, 18 und 19 der Handschrift DAM 01-27.1 (Sure 9:106-Ende, dann 19:1-67 und weiter)" in Markus Groß – Karl-Heinz Ohlig, eds., *Die Entstehung einer Weltreligion II. Von der koranischen Bewegung zum Frühislam* (Berlin: Verlag Hans Schiler, 2011): 311-402.
- Puin, Elisabeth, "Ein früher Koranpalimpsest aus Şan'ā' (DAM 01-27.1). Teil V: Die scriptio inferior auf den Blättern 14 and 15 sowie Auseinandersetzung mit den Thesen und der Edition des Koranpalimpsests von Behnam Sadeghi und Mohsen Goudarzi" in Markus Groß – Karl-Heinz Ohlig, eds., *Die Entstehung einer Weltreligion III. Die heilige Stadt Mekka – eine literarische Fiktion* (Berlin: Verlag Hans Schiler, 2014): 477-618.
- Robinson, Peter M.W. and Robert J. O'Hara, "Report on the Textual Criticism Challenge 1991", *Bryn Mawr Classical Review*, 1992, 3: 331-337.
- Sadan, Joseph, "Genizah and Genizah-like Practices in Islamic and Jewish Traditions", *Bibliotheca Orientalis*, 1986, 43: 36-58.
- Sadeghi, Benham and Uwe Bergmann, "The Codex of a Companion of the Prophet and the Qur'ān of the Prophet", *Arabica*, 2010, 57: 343-436.
- Sadeghi, Benham and Mohsen Goudarzi, "Şan'ā' 1 and the Origins of the Qur'ān", *Der Islam*, 2012, 87: 1-129.
- Sahle, Patrick, "What is a Scholarly Digital Edition?" in Matthew James Driscoll and Elena Pierazzo eds., *Digital Scholarly Editing. Theories and Practices* (Cambridge: Open Book Publishers, 2016): 19-39.
- Schmidt, Desmond and Robert Colomb, "A data structure for representing multi-version texts online", *International Journal of Human - Computer Studies*, 2009, 67 (6): 497-514.
- Schmidt, Desmond, "The role of markup in the digital humanities", *Historical social research*, 2012, 37 (3): 125-146.
- Segre, Cesare, *Semiotica filologica. Testo e modelli culturali* (Torino: Einaudi, 1979).
- Tomasi, Francesca, *Metodologie informatiche e discipline umanistiche* (Roma: Carocci editore, 2008).