Digital Editions of Medieval Texts

Contributed by Francesco Stella

Part 1: What’s in a name?

When we think about a digital edition, we must bear in mind what a Digital Edition is Not

‘A Digital edition is an online magazine or online newspaper delivered in electronic form which is formatted identically to the print version. Digital editions are often called digital facsimiles to underline the likeness to the print version. Digital editions have the benefit of reduced cost to the publisher and reader by avoiding the time and expense to print and deliver like a paper edition. This format is considered more environmentally friendly due to the reduction of paper and energy use. These editions also often feature interactive elements such as hyperlinks both within the publication itself and to other internet resources, searching and bookmarking, and can also incorporate multimedia such as video or animation to enhance articles themselves or for advertisements’.


This definition is quite useless to us (and in many particulars erroneous).

The scholar Patrick Sahle makes an important distinction

‘Digital scholarly editions are not just scholarly editions in digital media. I distinguish between digital and digitized. A digitized print edition is not a 'digital edition' in the strict sense used here. A digital edition cannot be printed without a loss of information and/or functionality. The digital edition is guided by a different paradigm. If the paradigm of an edition is limited to the two-dimensional space of the 'page' and to typographic means of information representation, than it's not a digital edition’.

(Patrick Sahle: Digitale Editionsformen, Zum Umgang mit der Überlieferung unter den Bedingungen des Medienwandels, 3 Bände, Norderstedt: Books on Demand 2013 (Schriften des Instituts für Dokumentologie und Editorik 7-9)

Wikipedia is effectively describes a digitized edition and not a digital edition.
First, we have to consider the question of what we mean by an edition, i.e. what are we trying to digitize? Be warned the meanings of “edition” in English are multiple, describing very different kinds of publication. Many of them are not very useful for our purposes.

1) An edition as a ‘Publication’. That is to say the first edition or reproduction of a text.

A publication of this sort can use any support such as stone, wax, papyrus, paper, wall, textiles, air, light, water and more.

2) A Documental or diplomatic edition. Such an edition reproduces a single existing source.

Conventionally, editions of all kinds are books.

3) A Critical edition. This is also called a reconstructive or philological or an analytic edition. This reconstructs the lost original or archetype of a text transmitted in multiple, different copies.

When such texts are encoded they may be online or a CD-Rom, i.e. digitized editions.
The concept of a ‘critical edition’ is key to digital editing. Many, if not most digital editions of medieval manuscripts are also critical editions. *English* Wikipedia doesn’t yet include an item “Critical edition” but it does have a not altogether satisfactory page for “Textual criticism”. So let us consider the definition of an *Edizione Critica* given on the Italian site.

Un pubblicazione del testo intesa a ricostruirne la forma originale, il più possibile rispondente alla volontà dell'autore, sulla base dello studio comparato (collazione) di ciascun passo dei diversi testimoni diretti e indiretti esistenti, siano essi manoscritti o testi a stampa. L’edizione si presenta perciò con un apparato critico che riporta le lezioni varianti.

(Cerquiglini: l’edizione deve rappresentare la mobilità del testo e non creare una stabilità mai esistita. Mordenti-Fiormonte: l’edizione critica è un prodotto della stampa.)

Translation:

‘A publication of the text intended to reconstruct the original form, according (so far as it is possible) to the author’s intention, through a comparative study (collation) of each step of the various direct and indirect witnesses, whether they are manuscripts or printed texts. The edition is therefore presented with a critical apparatus which presents the various readings.’

Cerquiglini: ‘The edition must represent the mobility of the text and not create a staticity that never existed.’

Mordenti-Fiormonte: ‘The critical edition is a product of the printing press’.
It is this kind of definition that informed that which was used in the *Parvum Lexicon Stemmatologicum* (2015).

‘A critical edition is an edition in which the text has been constituted on the basis of more than one source according to the genealogical principle. As it uses more than one source, in this respect it is an eclectic edition. What sets the critical edition apart is that it is based on a strict recension of the manuscript sources and the attempt to edit the archetypal text (possibly with some changes where the archetype is clearly faulty), and it is thus closely associated with the Lachmannian tradition of textual editing. Editors who do not agree with this reconstructive tradition usually refer to their traditions in other terms than critical’.


A critical edition, digital or otherwise, is therefore, the product of a specific, methodology and is terminologically distinct from those that do not observe that method.

*Scholarly Digital Edition* (Sahle, Robinson, Fiormonte): a digital publication made by scholars for an academic audience, presenting a text provided with scholarly data.

*Critical edition*: the publication of a text that is itself the result of a critical (philological, linguistic, cultural) process, which requires time, competences, and method. A reproduction of one of more witnesses of a text is not a critical edition.
Part 2: Digital versus paper editions

With the coming of digitization, the inadequacies of traditional, codex-format critical editions have become apparent.

Paul Maas: “the business of textual criticism is to produce a text as close as possible to the original”. In this tradition, criticism meant (and still means) methodological assessment of evidence following well-founded criteria.

Therefore, a text-edition should only be seen as fully critical if all interpretative decisions that led to the text are made as fully accessible and transparent as possible.

But: The cornerstone of true science is the principle of external replication ... The report on the research should contain sufficient information to enable peers to assess observations and to evaluate intellectual processes. That doesn’t happen in traditional editions.

(Vanhoutte)

SO...

Traditional critical editing, defined by the paper and print limitations of the codex format, is now considered by many to be inadequate for the expression and interpretation of complex, multi-layered or multi-text works.

(Marilyn Degan & Kathryn Sutherland, Text Editing, Print and the Digital World (2009), p.)
Is digital editing the answer to the problems inherent in the critical edition? All the way back in 1989, Bernard Cerquiglini certainly thought so.

Cerquiglini (1989!): ‘The screen, which technologists are forever improving and refining, is at the same time dialogic (it offers constant interaction between user and machine) and multidimensional (it allows the user to bring together, through the use of windows, data from unregulated sources). Making use of these two features, one can imagine a way of editing a medieval work, itself born of that collection of disparate units which is the codex, that would no longer be bound by the two-dimensional structure of the printed page: a diskette accepts varied masses of text which the reader looks at by bringing them up in different ways on the computer screen [...]’.


Yet for Cerquiglini such things remained a dream.

Rupert Pickens (1994): ‘even the prophet must admit that the appropriate ideal technology, in the form of compatible software, does not exist; indeed, he is forced into a kind of “nostalgia of the future” when ever greater technical elaborations and refinements will have produced the desired programs’.
Nevertheless, textual criticism has come increasingly to recognize the potential of digital editing and that texts are dissociable from their material supports.

‘For most purposes the fact of textual variance does not lead inevitably to the importance of variance... most reading and scholarly purposes require a stable text.’


‘The text does not have a material nature... the text is only and always an image, and any attempt to identify it with a material witness whatsoever, even an autograph original, is an attempt to conceal its unavoidable problematic nature.’

(C. Segre, Avviamento all’analisi del testo letterario (Torino 1985), p. 45)

The last major work on this subject is:
Robinson gives a good summary of the basic principles of preparing the material for a digital edition and so doing expressed the advantages of the digital edition.

It is worth reiterating the advantages of the computer analysis over traditional textual scholarship. In preparing the files for computer analysis the scholar makes no decisions about the significance or insignificance of a given reading. There is no weighting of evidence: indeed in this respect the system is counter-intuitive. When the blind test was being set up, I had assumed that some such weighting system would operate: that omissions would be accorded a greater weighting than mere variants, for example, and that sauts du même au même, because they are potentially polygenetic, would count for less than omissions where there was no textual stimulus to generate an eye-skip. To establish a hierarchy of evidential force seemed a necessary part of the procedure. This was an assumption inherited from traditional textual scholarship which proved to be unnecessary.

He proceeded to extol the benefits of the digital edition:

Lachmannian methodology is based on the identification and classification – the evaluation – of kinds of error. Thus, in traditional scholarship, assessing the significance of variants is a large part of the editorial process, and draws on the scholar’s expertise. Is a variant monogenetic or polygenetic? congiuntiva or separativa? an involuntary slip or a deliberate innovation? linguistically correct (or at least possible) or clearly erroneous? facilior or difficilior? All these kinds of discrimination count for nothing in the new context. For the computer analysis all that matters is accuracy – the accuracy with which the data is entered into the files.
Ultimately, Robinson concluded along these lines:

- Scholars often disagree in their evaluation of variants and the weight or significance to be attached to them. These disagreements often seem to be unresolvable.
- Such evaluations are also easily reversible; scholars can change their minds.
- Computer analysis gives us an answer, and that answer is independent of anyone’s opinion.
- the computer analysis dispenses with any need to evaluate them. It deals simply with the facts of the textual tradition, recorded as accurately as possible in the digital files, and produces its analysis accordingly.
- As our blind study established, the results of the computer analysis are uncannily close to those of a traditional Lachmannian analysis for the groups at lower levels of the stemma. Only the archetype eludes the computer analysis, since what the computer produces is not strictly speaking a tree but a phylogram: the proof of an archetype remains as elusive as ever.
- Any form of selection or weighting of the data involves the operation of subjective human judgment, meaning different people will produce different results, and the disagreement between them will be unresolvable.
Not everyone, however, agreed. Ben Saleman pointed out that the process of selecting data to be encoded remained a selective process:

‘Often the status of variants, the textual differences, seems to have become unimportant to them. They simply consider each textual difference as an objective, easily observable (objective) fact, although they sometimes exclude small or unimportant variants for unexplained reasons. They gather these objective facts and introduce them in statistical-mathematical software which builds, in an objective way, a tree out of them. But is such a tree a chain or a stemma, a text-historical tree?

In inductive research, the objective facts must be related to the goal of the research. If I want, for example to predict the weather, I can gather all kinds of objective facts in and around my house: paperclips, stones, papers, etc. It is obvious that I will not be able to predict the weather with these facts, even though they are objective.’

(Ben Saleman, Building stemmas with the computer in a cladistic, neo-Lachmannian way (Nijmegen: 2000), p.)

In 2013, Robinson issued a retraction

‘There has been a great deal of rhetoric, some of it from myself, in the last decades about how scholarly editions and editing have been fundamentally changed by the digital turn. So let me say it plainly. I don’t think there has been any such change. A scholarly edition is still, as it has been for centuries, an argument about a text. The fundamental players in this argument are still documents, works, and the editors’ interpretation of them. The editor is the editor, and not a “facilitator”. There are still many more readers than editors, and most readers do not want to be editors.’

(P. Robinson, ‘What Digital Humanists don’t know about Scholarly Editing, and Scholarly Editors don’t know about the Digital World’, paper presented at the conference Social, Digital, Scholarly Editing, University of Saskatchewan, July 11, 2013-July 13, 2013, accessible online under the title “Why digital humanists should get out of textual scholarship”.

Francesco Stella—Digital Editions of Medieval Texts

Digital Editing of Medieval Manuscripts - Intellectual Output 1: Resources for Editing Medieval Texts (Paleography, Codicology, Philology)
Part 3: Producing a Digital Edition

The options available to the digital editor are considerable:

**Contextual editions** combining primary texts and secondary contextual material.
(for example: http://www.woolfonline.com/)

**Genetic editions** where the text is presented as having been constructed from a collection of other documents.
(For example, Wittgenstein’s Nachlass: http://wab.uib.no/wab_BEE.page)

**Philologic** establishment of the scholarly “best” text
(For example, the editions produced by the Institute for Textual Scholarship and Electronic Editing, e.g. the editions of Chaucer: http://www.birmingham.ac.uk/research/activity/itsee/publications/canterbury-tale.aspx)

**Hybrid** (archives+forthcoming edition, ex. Henrik Ibsen’s Writings)

The editorial process must be user oriented and it has 4 steps:

1. content gathering and processing.
2. generation of intermediary representation by means of encoding.
3. transformation of these representation by means of algorithmic mechanism.
4. output of presentation by means of an ultimate transformation.

Creators of digital editions must first consider the constraints that are imposed by digital editions.

- The Availability of all sources
- The verifiability of the philological choices as a condition of scientific reliability
- The visualization of textual multiplicity such as mobile archetypes, the chronological dimension and geographic versions

Remember, as R.B.C. Huygens pointed out:

“even if you try to reconstruct the oldest attainable stage of the manuscript tradition, which should be your aim, you must nevertheless be aware of the fact that ... the original itself played much less important a role, if any at all, than its often defective descendants”
The creator of a digital edition, will also be confronted with the choice as to whether they will include any further features. The following are just some of the possibilities.

- Dimension (unlimited) > reproduction and/or integral transcription of the witnesses > edition-archive
- Multimediality > audio
- Hypertextuality > additional informations set or versions comparison linked to the main text
- Interoperability > fast corrections, scholarly discussion in “real time”, “transferability”
- Improvement > instability, hardness of quotations
- Multiple visualization: able to represent the diachronical mobility of a text
- Duration?
- Costs
Part 4: Before the Edition

1. Transcription

Reading (the Colour Multispectral Scanner)

Even when a manuscript appears to be illegible, digital tools can still come to the scholar’s aid.

Through spectroscopic scanning you can separate the different layers of writing by marking each by a different colour. In this way, even the lower writing, quite imperceptible to the naked eye, can be read.
The images below show the multispectral imaging of a palimpsest from Firenze, Biblioteca Medicea Laurenziana, MS. Plut. 87.21, fols. 4v-5r.

This particular text is bis rescriptus, so there are three layers of writing):

1. IX century: Historia Philothea of Teodoreto of Cirrus.
2. XI century: Iliad.
3. XIII century: several texts, among them some Aristotle’s pieces of work.
The results reveal legible Greek text beneath the upper layers:
Such a visualisation technique was used in October 2014 to reveal the lost text on an 800-year-old manuscript of Magna Carta. (See: http://www.dailymail.co.uk/sciencetech/article-2784884/Magna-Carta-manuscript-reveals-secrets-Scans-uncover-lost-text-hidden-damaged-section-800-year-old-parchment.html )
Even very rough and ready tools such as xnview can reveal usable text. Xnview is little more than a multimedia viewer that allows you to adjust brightness and contrast. In the case of the manuscript below, this simple tool made the illegible, legible once more.

Mino da Colle, Firenze
Bibliotheca Nazionale,
MS. n.a. 285, fol.
From such efforts come the raw materials for digital palaeography projects. Here is an exposition on the methodology of digital palaeography taken from the dissertation of Arianna Ciula.

The parameters guiding the segmentation [categorization] are attuned to the typology to be segmented, that is to say, according to, what is called, an ‘x’ projection of a character. The letters are ideally divided into three groups depending on the expected areas of ink distribution. Particularly ‘unimodal’ characters (such as ‘I’ or ‘l’) shall be part of the first group; those which are especially ‘bimodal’ the second (for example ‘b’ and the linked ‘st’); those which are ‘trimodal’ the third group (for example ‘m’, the linked ‘sti’). Characters that have a histogram that is expected to be variable according to the style of the writer, as is the case, for example, with the letter g, come from time to time [and] are segmented by selecting the corresponding ‘modality’.
For examples of sites that are concerned with digital palaeography, please see:

- Werkzeug zur paläographischen Dokumentation von Handschriften—Based at the University of Köln (http://www.ceec.uni-koeln.de/projekte/CEEC/tools/paleography/dokumentation-palle.htm)
- Olaf Puta’s Abbreviationes—an online database of Latin abbreviations created by the titular Olaf Puta (http://olafpluta.net/software/software.html)
- Encoding Boccaccio’s Zibaldone—an online edition of Boccaccio’s autograph edition of Zibaldone (http://rmcisadu.let.uniroma1.it/boccaccio/)
- The Medieval Palaeographic Scale—an effort to date manuscripts according to the handwriting that they contain based at the Huygens Institute for the History of the Netherlands (https://www.huygens.knaw.nl/mps-de-medieval-palaeographic-scale-ter-datering-en-lokalisering-van-middeleeuwse-handschriften/?lang=en)
- DIGIPAL—A resource for the study of medieval handwriting, particularly that used in England between 1000 and 1100 (http://www.digipal.eu/)
- VHMML—A digitisation of manuscripts from hundreds of libraries sponsored by the St John’s University in Minnesota (https://www.vhmml.org/)
- Ad Fontes—A learning tool from the University of Zürich (http://www.adfontes.uzh.ch/1000.php)
Part 5: After Transcription: Encoding?

Encoding is the conversion of information into another form of representation for communication in a medium.

First the edition’s creator must ask what can or should be encoded? The answer to this is that a text can be encoded, but the whole edition cannot. “What can be encoded is just a limited part of the complex piece of work we call “digital edition”” (Robinson)

 Encoding should be done to a standard specification and TEI is among the best and most flexible tools for encoding digital editions. TEI allows you to mark-up both the primary source and may be used to incorporate critical apparatus (though it is not necessary to do so).

The idea of a standardised mark-up has not always been accepted. At the founding conference of TEI (1987), Ian Lancashire opined that:

“No one editorial standard has emerged after 3,000 years of work, and I am skeptical that one will emerge. A scholar should have the right to publish texts encoded in the way he or she believes best suits the purpose of the analysis.”

However, standardisation has numerous benefits, not least that it makes the end result simpler for other scholars to interpret in its mark-up form and the editorial decisions more transparent. TEI is a particularly good standard. It is flexible, coming with a large range of tags and with highly detailed instructions on how to use it. (See the TEI wiki: https://wiki.tei-c.org/index.php/Main_Page)
TEI also has the advantage that it comes with a full range of standardised mark-up tags with which to describe a manuscript, such as:

- Suspension: first letter(s) of the word or phrase, omitting the remainder.
- Contraction: omits some letter(s) in the middle.
- Brevigraph: comprises a special symbol or mark.
- Superscription: writing above the line.
- Acronym: the initial letters of the words of a phrase.
- Title: abbreviation for a title of address (Dr, Ms, Mr, ...)
- Organization: name of an organization.
- Geographic: for a geographic name.

So For Example

<choice>
<abbr>SPQR</abbr>
<expan>senatus populusque romanorum</expan>
</choice>

<expan>senatus populusque romanorum</expan>
</choice>
It is possible to encode multiple readings:

```xml
<app>
<lem wit="#El #Hg">Experience</lem>
<rdg wit="#La" type="substantive">Experiment</rdg>
<rdg wit="#Ra2" type="substantive">Eryment</rdg>
</app>

Corrections
And of so parfit wis a <app>
<rdg wit="#Hg">wight</rdg>
<rdg wit="#Ln #Ry2 #Ld">
<corr resp="#ETD">wright</corr>
</rdg>
<rdg wit="#Gg">
<corr resp="#mp">wyf</corr>
</rdg>
</app>

And more
It is worth giving some quick definitions of different varieties of encoding your text before we continue:

**Location-referenced method**— The apparatus uses like numbers or other canonical reference scheme referenced in a base text. This is useful for encoding the apparatus in printed text.

**Double End Point Attachment**— The apparatus indicates the precise locations of the beginning of each lemma relative to a base text. It differs from the location referenced method in which only the larger span of text containing the lemma is indicated. In this method, it unambiguously matches each variant reading against its lemma.

**Parallel segmentation**— Alternate readings of a passage are given parallel in the text: no notion of a base text is necessary. All the texts compared are divided into matching segments all synchronised with one another.

**Stand Off**— This is a kind of mark-up that resides in a location different to that of the data being described by it. It is the opposite of the three kinds of mark-up described above, which mix markup with the text being described.

For more detail, see the relevant pages on the TEI wiki (https://wiki.tei-c.org).
It would be wrong to suggest that TEI mark-up is a perfect tool. It does have a number of drawbacks. The key text for thinking about the problems of mark-up is:


In Abstract:

Embedded generalized mark-up, as applied by digital humanists to the recording and studying of our textual cultural heritage, suffers from a number of serious technical drawbacks. As a result of its evolution from early printer control languages, generalized mark-up can only express a document’s ‘logical’ structure via a repertoire of permissible printed format structures. In addition to the well-researched overlap problem, the embedding of mark-up codes into texts that never had them when written leads to a number of further difficulties: the inclusion of potentially obsolescent technical and subjective information into texts that are supposed to be archivable for the long term, the manual encoding of information that could be better computed automatically, and the obscuring of the text by highly complex technical data.
For example, it is sometimes recommended that 'stand-off' apparatus is used instead of parallel segmentation to avoid problems with scaling:

Scaling is a problem with methods of indicating textual variance, but in parallel segmentation this is exacerbated because as the number of witnesses increases, the likelihood of needing to reformulate the reading boundaries, never mind the difficulty in reading or understanding such encodings. This may be a problem not only when looking at a single text with many witnesses, where variation in the structure may be extremely difficult to represent where conflicts occur which disrupt this very basic structure (for example imagine a set of witnesses where some have lines in linegroups, some just lines, some paragraphs, some paragraphs in divisions, but all with the same underlying text). But also where parallel segmentation is being used to record divergent interpretations of these individual witnesses by many editors (for distributed cooperative editions generated from many editorial views of a text). A plausible recommendation is to use a form of stand-off apparatus for such editions rather than parallel segmentation. And while some of the current methods can be used in a stand-off method, they should be updated to reflect current usage of URI pointers.

for a more detailed account of this procedure, see "Intellectual Output" N. II (ed. M. Burghart)
Part 6: After transcription: Collating, Some Software?

The Versioning Machine

The ‘versioning machine is a stand-alone editing environment and interface for displaying multiple versions of text encoded according to the TEI. It uses html, css, javascript, xslt, xml and encodes every witness in a single structure. Repeated lines are not encoded again.

It produces editions of “historical texts” (Tanselle): “each version of the work being a witness to a textual moment” – It can produce an image of progression or snapshots. In its laboratory like setting, links with facsimiles can be embedded in it. Insertions and erasures may be visualised in colour.
Another alternative is Collate, a tool designed by Peter Robinson based on his study of forty-two manuscripts of the two Eddic Old Norse poems Gróugaldr and Fjölsvinnsmál.

Collate works interactively with the collation being written to a window as the scholar watches. The scholar may intervene at any point to alter the collation, using either of the tools "Set Variant" or "Regularise". "Set Variant" allows the scholar to over-rule the collation offered by Collate and impose his own collation, even writing a variant that does not appear in the sources into the collation. Collate includes a particularly powerful regularisation facility, derived from my struggles with the highly individual orthographies and spelling systems of Icelandic scribes. "Regularise" enables the scholar to intervene to regularise any word or phrase in any source at any point. The regularisation can be set for a particular word at every point in every source, or for that word only at that place in that source, or various other combinations. Collate will record all variants set and every regularisation made and remember them next time it runs. The scholar can adjust the collation in other ways, switching the base text, suppressing agreements with the base text and collating punctuation tokens separately. For more, see (http://www.hd.uib.no/humdata/2-91/robin.htm)

Collate X

The successor to Collate is the Java software tool Collate X. It was developed by several partner institutions under the umbrella of the European initiative 'Interedition'.
**An Example, The Donne Variorum:**

This site, which is the online counterpart of the *Variorum Edition of the Poetry of John Donne* (8 vols. Indiana UP, 1995-), compares versions of a same texts, the writings of John Donne, saved in txt format, each one provided with:

1. **Identiline:** filename, source name, folio and/or page locations in the source, initials of the transcriber, date of the transcription, etc.

2. **Header line(s) -** contains the title

3. **Body lines -** one line for each line of poetry

4. **Subscription line -** subscription or attribution

5. **Information line -** a noteworthy features may be recorded.
Outputs of collation:

- **Exact Matches** - The program suppresses the word in the variant line, leaving a blank space in the output. This reduces clutter in the collation and allows you to focus on the variants.

- **Alterations** - If the word in the variant line differs in any detail from that in the base line, the variant will be situated directly under the base word and printed out in full.

- **Insertions** - If a word appears in the variant line that has no counterpart in the base line, it will be interpreted as an insertion, enclosed in braces ({insertion}), and printed out at the appropriate place in the collation.

- **Omissions** - If a word that appears in the base line does not appear in the variant line, the omitted word is enclosed in angle brackets (<omission>) and printed directly under the base word.
Francesco Stella—Digital Editions of Medieval Texts

Part 7—Step 4: establishing connections among the witnesses

Stemmatics?

See:

Divina Commedia edition of P. Shaw–P. Robinson, SISMEL 2010 (www.sismelfirenze.it)

A traditional Stemma by Petrocchi

An Unrooted philogenic stemma

and: 14 witnesses
ex: and: 28 witnesses
and eek: 11 witnesses
also and: 1 witnesses
Robinson: apparato a comparsa

This system can highlight the variants of a single word...

...every word of a line...

...all the subvariants of every word in a line...

Extended Apparatus

Line 10

Veramente: Ash, Ham, LauSC, Mart, Rh, Tr, Urb
quantum del: Ash
quantum de: Ham
quantum id el: LauSC, Mart, Rh, Tr, Urb, FS, PET
regno: Ash, Ham, LauSC, Mart, Rh, Tr, Urb
santo: Ash, Ham, LauSC, Mart, Rh, Urb

Robinson 2002—Paradiso 1:10

(extended)

...the number of witnesses for this line...

...and the types of the line’s forms with their sigla.

Coloured Differences

Selected Witnesses
Stemmatology
The Stemmaweb team is experimenting a new, more philological, method to select variants by giving to each category a different weight and to visualize their distribution among the witnesses.

‘Given a collection of imperfect copies of a textual document, the aim of stemmatology is to reconstruct the history of the text, indicating for each variant the source text from which it was copied. The project develops theory and methods for computer-assisted stemmatology, and evaluates the accuracy of such methods in simulated and real data-sets.

Stemmatology lies at the intersection of several scientific disciplines. On one hand, it is associated with humanities where texts are used as sources, and on the other hand, to mathematics, statistics, and computer science, and finally, to evolutionary biology and cladistics, the study of evolution and speciation. The aim of traditional stemmatology — or textual criticism — has been to infer the original content of a textual source based on a number of different versions. Modern computer-assisted stemmatology has proven to be an
Part 8: Visualizing the edition

A. Archive Editions

Fontes Civitatis Ratisponensis

One of the earliest digital editions of archive documents, started in 1996 and closed in 2006.

(http://bhgw20.kfunigraz.ac.at/, accessed 27th May 2017)

Other pioneering projects achieved by the Center for Computing in the Humanities (now Dept of Digital Humanities), King’s College, London.

Anglo Saxon Charters

(http://www.aschart.kcl.ac.uk/index.html, accessed 27th May 2017)

Fine Rolls

Codice Diplomatico Lombardo
two of the earliest Italian digital projects on medieval archives.

projects of editions and toolboxes promoted by the Ecole nationale des Chartes, Paris.

Liber Matriculae

Éditions en ligne de l’École des chartes (Élec)

Diple, modular methodology and tools for heterogeneous TEI corpora
Part 8b: Literary editions

Beowulf: one of the earliest digital editions ever, first published in dvd, then moved to the web, which had to be converted three times in different platforms and codes for avoiding access restrictions or interoperability obstacles due to proprietary parts of the source or obsolescence.

Ebeo 4.0
(http://ebeowulf.uky.edu/ebeo4.0/start.html, accessed, 27th May 2017)
Provisionality

The third edition of Electronic Beowulf was an html application on DVD that used a Java applet and JavaScript. When first published, major internet browsers could run the html application on PCs and Macs. However, security problems with Java in Summer 2013 forced all major browsers to disable unsigned Java applets compiled with earlier versions of Java. As it was compiled in 2011, Electronic Beowulf 3.0 was then disabled.

To fix the problem in 2013 we compiled a new signed applet with Mac and PC installers that moved Electronic Beowulf 3.1 from the read-only DVD to the owner's hard disk. By 2014, however, this solution was beginning to have problems, because Java kept changing its security protocols. To solve the problem, we decided to stop using Java altogether and to re-engineer Electronic Beowulf using only JavaScript. At the 49th International Congress on Medieval Studies in May 2014, Emil Iacob explained the technological issues and Andrew Prescott announced in a plenary lecture celebrating the twenty-first anniversary of Electronic Beowulf that the fourth edition would be going online in 2015.
Editionsproben
One of the best screen partitions for texts, with multiple witnesses: a quadrant with 1 base (or critically reconstructed text), 2 apparatuses with list of the manuscripts, 3 texts of one selected manuscript, 4 its image.

Updating 2015
2015: second version of the Parzival, updated for the same purposes as the Beowulf. A sign of the digital edition’s instability and provisionality.

Fassungen
(See: http://www.parzival.unibe.ch/home.html, accessed 27th May 2017)
Piers Plowman Electronic Archive

Here again a sequence of the different versions of the digital Pier the Plowman, from 1994 to 2014, and some samples of its views.

New 2011 Piers

Last 2014 Piers

Views

Technicalities

(See: http://piers.chass.ncsu.edu/, accessed 27th May 2017)
The DBR (Data-base of Rhythmi) is the Visual Basic input software the research team uses for insert the data about texts, musics, language, manuscripts. It is extensible and adaptable to whatever content, by only renaming the records fields, and has already been used for other projects (such as the Carolingian texts – before and after correction – of Turin Hagiographic Collection edited by M. Goulet).

This is the interface where to put texts and apparatus, strophe by strophe. No need encoding, since it is a database.
DBR Language Filing

The Language filing plans to record on one side any phonetic, morphological, syntactic and semantic difference with classical Latin, on the other side the statistical figures about numbers of names, adjectives, verbs and their reciprocal distance, giving a measure of the closeness to the so-called proto-romance.

Public Version (CRM): Index

General Information in the Text

The last screenshot is the opening view of the public version (cd.rom or website): list of included texts, general data, critical edition (reconstructed according to Lachmannian criteria or not)

Analysis of the versification: symbolic representation of the structure both of the line and of the strophe in two different systems (the traditional Norberg and D’Angelo, more suited to medieval treatises)
The table of correspondence between Noberg's and D’Angelo’s system
The apparatuses of variants, loci paralleli divide into sources, contemporary passages and afterlife.

Synoptic edition of texts with multiple versions, stemma with clickable sigla of every single manuscript

Clickable Stemma Codicum
Francesco Stella—Digital Editions of Medieval Texts

Manuscripts

The core of the edition: the reproduction of the MSS with the relevant transcriptions, readable both with or without diacritical signs MSS Images and Transcriptions

Music transcripts and records
Francesco Stella — Digital Editions of Medieval Texts

Tables of interpretations of the alphanumeric code created by Sam Barrett (Cambridge) and used for the computer-readable transcription of the musical neumes.
Francesco Stella—Digital Editions of Medieval Texts

Linguistic Archives
the core of the edition: the reproduction of the MSS with the relevant transcriptions, readable both with or without diacritical signs

Quintuple Display
So the edition can display 6 versions of the same “song”: 1 visualized manuscripts, 2 textual transcription, 3 musical transcription by hand, 4 musical transcription on stave, 5 musical conversion into alphanumeric code, 6 reconstructed edition.

palaeographical information about the manuscripts, which can also be leafed through one by one for didactic or scholarly purposes.

Digital Editing of Medieval Manuscripts - Intellectual Output 1: Resources for Editing Medieval Texts (Paleography, Codicology, Philology)

Co-funded by the Erasmus+ Programme of the European Union
Search: search engine which can work by crossing every field of the database, so selecting texts or transcription with a certain metrical structure or melodic outline or linguistic phenomenon etc.

Searchsample: Italy+viii century+consonant exchanges
You can also produce lexical concordanced of the tokens or statistics of the most frequent words or roots or endings both searching though the editions and the single transcriptions.

Sortable linguistic statistics
An interactive table of statistics is ready, collecting the figures about numbers and distances of names, verbs and adjectives.
To finish, some advice:

- The creator should aim to make the project fit the resources and the available time, i.e. it must be sustainable
- Don’t trust methods used in demos, unachieved works, samples.
- Create the added value of your digital edition
- Don’t be content with simple transcriptions of single witnesses just because it’s easy and quick and don’t justify what is easy and quick as if it was the best option.
- Edit every text as if it was your text.

A Typological Divide:

- Diplomatic Edition (documents or single-witness text or mobile texts): limited number of features > TEI, automatic editors.
- Philological edition (with reconstruction of the archetype or of the original, analysis of the relations between the witnesses, and linguistic analysis): unlimited number of features and witnesses > creative common DB or open interface linking multiple plain-text or multimedia files
Constitutio textus: Possible Scenarios

1) Refinement of variants analyzers which lead to a critically founded stemma (ex. Stemmaweb) or

2) Renunciation of a digital processing of the variants, then

3) Transformation of the concept itself of c.e.: giving up a digital edition which is also really critical (Cerquiglini) and flourishing of “documental” d.e. of single mss. devoided of authorial representativeness. We accept to never know what the author wrote

Publishing

a) Development of a TEI critical edition scheme which is accepted by the most exacting scholarship (in progress, very slowly).

b) Creation of tools automatizing the encoding procedures (possible only after the new TEI scheme is developed).

c) Renunciation of the embedded encoding and transfer to a stand-off encoding

In the meanwhile

Using computing tools to visualize and analyze editions created by hand and to make sources available

Danger

Division between “high-level” reconstructive paper philology, more correct on the methodological point of view but without technical possibility to preserve the evidences and “low-level” semi-critical transcriptions of single real sources on digital supports