i|d|e

ROMAN BLEIER   SEAN M. WINSLOW (EDS.)

VERSIONING CULTURAL OBJECTS: DIGITAL APPROACHES

herausgegeben von:

Bernhard Assmann	Roman Bleier
Alexander Czmiel	Stefan Dumont
Oliver Duntze	Franz Fischer
Christiane Fritze	Ulrike Henny-Krahmer
Frederike Neuber	Malte Rehbein
Patrick Sahle	Torsten Schaßan
Markus Schnöpf	Martina Scholger
Philipp Steinkrüger	Georg Vogeler

Band 13

Contents

Electronic Versioning

Appendix

Preface

The present volume, *Versioning Cultural Objects: Digital Approaches*, is a collection of selected essays that were first presented and discussed at a symposium at An Foras Feasa, The Research Institute for the Humanities at Maynooth University, in December 2016. The idea of the volume is to start a discussion about the different types of versions we are dealing with in the digital humanities (texts, objects, analogue, and digital resources) across disciplines.

Graz, December 2019, the editors

Introduction

Introduction: Versions of Cultural Objects

Roman Bleier, Sean M. Winslow

Abstract

The *version* of a cultural object is identified, defined, articulated, and analysed through diverse mechanisms in different fields of research. The study of versions allows for the investigation of the creative processes behind the conception of the object, a closer inspection of the socio-political contexts that affect it, and may even provide the means to investigate the object's provenance and circulation. At a symposium at Maynooth University, scholars from different research areas exchanged ideas about different forms of media, including text, image, and sound, to work towards an understanding of the term *versioning* in the broadest sense. While the understanding of versions and related terminology differs between disciplines, a cross-disciplinary dialogue will highlight the range and depth of existing studies and provide an interdisciplinary understanding of the term versioning which will be useful for a more holistic conceptualisation. The present volume tries to contribute to this dialogue by providing eight peer-

reviewed articles resulting from the discussion and presentations held at Maynooth University.

The breadth and applicability of the concept of a *version* is at the core of this volume. Questions like: *Can the word* version *be applied uniformly across disciplines? Does the meaning of the word change?* drove the editorial decisions in bringing together the various participants in the original symposium in Maynooth which was the beginning of this volume. The range of the answers presented here underline the striking multivariance of the term, and the way that different humanities researchers are using it, from music to genetic criticism to versioning as it is understood in the management of shared code databases. By choosing these articles, we hope that we can offer not only a sense of the range of the field, but invite the reader to think about the many facets that have to be considered in order to fully understand the semantic lifting done whenever the word version is encountered, and how we might begin to form a shared understanding of the fullness of the term, but also where it needs more support and specificity.

1 The genesis of this volume

This volume had its genesis in the work of An Foras Feasa during 2015 and early 2016, The Research Institute for the Humanities at Maynooth University, then headed by Susan Schreibman. Roman worked on the redesign and release of the *Versioning Machine 5.0*, a publication framework for the display and visual analysis of multiple versions of a text. Vinayak worked on a theoretical framework to capture electronic metadata of visual resources (see Das Gupta); it was in that setting that the question arose of how to record reproductions and the context they were produced in. In order to foster an interdisciplinary discussion about the

topic, they organized a symposium as a platform to present and discuss the various disciplinary approaches. In addition to the presentation of papers, the participants worked in groups to examine related terminology. This cross-disciplinary exchange can be seen in the finished chapters.

2 Why was the term *versioning* used?

The term versioning is more frequently used in the context of software versioning and electronic version control. The *Versioning Machine*, developed by Schreibman et al. in the early 2000s (launched in 2002), introduced the term in the sense of *exploring variation between textual versions of a work* into the digital humanities community (see Schreibman, "Re-Envisioning"; Schreibman et al., "The Versioning Machine"). With the *Versioning Machine*, Susan Schreibman investigated the composition process of Thomas MacGreevy's poetry by comparison and parallel reading of various versions of the poems.

   Taking Schreibman's work as a point of departure, and the attendant realization that versioning means different things to different disciplines and to different practitioners, the edited articles in this volume illustrate the range and depth of existing studies of versions and will (we hope) provide a first step towards a platform for an interdisciplinary discussion and understanding of the concept. The volume engages with versioning in the digital humanities in three primary areas: the conceptualisation of versions in different humanities disciplines, the methods involved in the electronic modelling of versions of cultural objects, and the representations of digital versions. Individual articles may cover one or more of these areas in varying depth. Appropriately enough for a book on versioning, our volume opens with Elisa Nury's dissection of the meaning of *variant reading* in textual scholarship. She asks whether the

concept of "variance" means the same thing in different disciplines, emphasizing the importance of contextualisation of the term and presents an implementation of a digital representation of a *reading*, which is a first step to conceptualise variant reading, using the CollateX JSON data format.

3 Textual versions and digital editing

The advent of the digital age has led to a profusion of digital versions of documents, but problems in dealing with versions are hardly new: palaeographers had to deal with different versions of scripts, numismatists with versions of coins, archaeologists for instance with marble versions of Greek bronze statues or motives on Greek red-figure pottery, textual scholars with versions of written sources, art historians with different versions of artworks. The methodologies developed in a pre-digital context still have validity today and many scholarly discussions have continued and are being adapted in the digital scholarly context. So what does it mean when we, as digital humanists, talk about versions? Where do traditional approaches of the pre-digital age end, and what do new, digital approaches entail?

For instance, digital textual editing discussions about versions go in different directions: in stemmatology and copy-text editing, an editor has to establish which variation between different manuscript witnesses to "trust" in order to establish a "safe text" that comes as close as possible to an author's original work. Editors following the genetic editing approach try to untangle the various layers of revisions and changes made to a manuscript over time. While genetic editing was the exception in print, the flexibility of the digital medium to represent different layers of a text has led to a substantial increase in the

development of such editions (see Pierazzo, "Digital Documentary Editions"). One of the central characteristics of digital scholarly editing is the separation of data and presentation. The data is usually represented using the standard of the Text Encoding Initiative (TEI) which allows the modelling of versions of a text in concordance with the traditional editing approaches (see Burghart; Pierazzo, "Facsimile").

Three chapters in the volume, by Martina Scholger, Richard Breen, and Christian Thomas present case studies of digital scholarly editing projects that investigate different kinds of versions and variance.

Martina SCHOLGER's chapter, "Pieces of a Bigger Puzzle," explores her work on a digital scholarly edition of the notebooks of Hartmut Skerbisch, an Austrian visual artist. The notebooks contain a network of references to music, literature, and other visual art works as well as numerous sketches, constructional drawings, and diagrams of Skerbisch's installations in various stages of conceptual planning. Lacking the finished installations, Scholger uses a genetic criticism approach to uncover and identify the various versions which are the result of the artist's creative process and to examine the relationship between sketch and visitors' reports of the final installation in relation to the genesis of his artistic work.

RICHARD BREEN explores the transmission of the many variants of "The Unfortunate Rake." We might wonder what "St. James Infirmary Blues" has in common with "Streets of Laredo," a nineteenth-century cowboy song, or what either has to do with an Irish folk song. To explore and visually show the motivic similarities, Breen uses the *Versioning Machine* and StoryMapJS to map and narrate the distribution of the song variations across the globe. Versions in this case are similar motives that developed across variations of the song, uniting a seemingly-disparate corpus in one family network.

Like Scholger, Christian Thomas seeks to reconstruct a missing event, in this case, Alexander von Humbolt's Kosmos-lectures from Humbolt's fragmentary manuscripts, lecture notes taken by attendees, and related documents. These fragmentary reports about the lectures can be viewed as witnesses or versions that can enrich and complement our knowledge of the lectures and its contents, but they may also present conflicting information and narratives. The question is how to deal with such a rich and diverse number of primary sources, especially if–like in the case of the lecture notes–their authorship and origin is not always clear.

4 Considering other representational forms as versions

Historians have for many decades made editions where *regest*s, short abstracts listing the main information about a text, have been used for extracting and summarising information important for historical research. Focusing on content–rather than wording–allows the creation of versions of texts enriched by external information (in the form of RDF) in order to to find connections and support advanced search functionalities goverened by a conceptual model (see Vogeler). Consequently, as representations of the information layer, abstracts, regests, or metadata should be considered as expressing a version of the same, each supplementing or replacing other possible versions based upon project and disciplinary needs. In this volume, Georg VOGELER uses the example of medieval charters to discuss copies and what other kind of versions were added in the digital world: transcriptions and reproductions of a charter in print and digital form, archival and scholarly descriptions, and metadata become part of his model. He suggests a graph-based data model with RDF that allows a more

flexible and suitable approach than current XML or relational database solutions.

This focus on the information layer also leads one to think about the various trajectories in the production and commodification of an object, and the meanings and values associated throughout these histories, which can be referred to as *object biographies* (see Kopytoff). Treating the history of objects as a version of what they are, fully in parallel with their content, reminds us that objects and texts, as they come down to us today, may not only exist in different versions (as of, say, a painting), but are usually different versions of themselves, having undergone changes, whether physical or in terms of their recontextualization, which affect our interpretation, and are themselves, in effect, variant readings of the object. That these changes should have a temporal aspect which needs to be considered is no surprise, and ATHANASIOS VELIOS and NICHOLAS PICKWOAD's contribution to this volume, presenting CIDOCCRM events for reconstructing the history of binding structures, attempts to address the need to formally document this temporal aspect in digital codicology.

5 Electronic texts and version control

The book concludes with two papers discussing principles of the versioning of electronic documents, comparing versions of electronic documents, and problems when trying to collaboratively work with documents in an online environment. Specifically-electronic considerations for editing should be taken into account, as in GIOELE BARABUCCI's exploration of different abstraction levels of electronic documents which can be described by their content, model, variants, and physical embodiment. The paper describes the problem and presents a formal solution in the CMV+P model. The implementation of this model would enable a

user (human or computer) to precisely describe and communicate the type of version of an electronic document the user is interested in. Practical applications include document comparison tools which could operate on a CMV+P based model to compare only the levels of primary interest.

Metacontextual issues around project management and collaboration are considered by Martina Bürgermeister, who discusses the importance of versioning control systems for digital collaboratory research environments. She critically analyses collaborative projects such as Annotated Books Online, Monasterium.net and Wikipedia by exploring how collaboration and versioning control is implemented by these organizations. She concludes that existing collaboratories do not satisfy the needs of humanities research, and suggests conceptual models which will help us to classify the various types of changes happening in electronic documents during collaborative work and the relationships between them.

6 Concluding remarks

In a way, a single thread connects Nury's opening article, which starts the volume with a solid grounding in the text critical concept of versions, to Bürgermeister's closing article, which deals with versioning metadata from current development practices. We proceed from the big, basic question of what a version is, through case studies, to domain-specific formal systems for representing knowledge (discussed by Vogeler, Velios and Pickwoad), down to a narrowed and focused exploration of the actual codepoints that represent word information (Barabucci). Another thread might go from Nury's analogue context, through case studies by Breen, Thomas, and Scholger–which all present material which could be represented in an analogue edition,

but where digital methods help to present the complexity of the data more clearly than was possible in print–to the purely-digital representations enabled by graph data and the digital structure of word data itself, and the data about data that is collected by a versioning system. Yet another thread would wind in a convoluted and hopelessly knotted fashion, detouring for all the similarities among the articles. As an example, in both Scholger's work on artists notes and Thomas' work on the Kosmos-lectures, we lack a direct and authoritative version of the "main event" (the installation for Skerbisch and the lectures for von Humboldt), leading to a reconstructed *ur*-version which is itself unstable and subject to variance in interpretation. Here, we see techniques developed for critical textual editing applied to the reconstruction of performance. Vogeler's work on charters highlights similar issues for drafts, as the final, legal version of a charter can be preceded by non-legal drafts, and followed by promulgations and re-issues that are separate legal acts of the same basal charter text. Here, the versions speak to both the textual development of the charter as well as the various instances of its legal effectuation. We hope that these examples will encourage people to give thought to how the concept of versioning changes and with what kind of new versions we are dealing in a digital context.

Bibliography

Burghart, Marjorie. "Textual variants." *Digital Editing of Medieval Texts: A Textbook*, edited by Marjorie Burghart. 2017. www.digitalmanuscripts.eu/digital-editing-of-medieval-texts-a-textbook/. Accessed 21 Aug. 2019.

Das Gupta, Vinayak. "Albums in the attic." *An investigation of photographic metadata*. Studia Universitatis Babes-Bolyai 62. 2017, pp. 57-74.

Kopytoff, Igor. "The cultural biography of things: commoditization as process." *The Social Life of Things: Commodities in Cultural Perspective*, edited by Arjun Appadurai, Cambridge University Press, 1986, pp. 64-91.

Schreibman, Susan, et al. "The Versioning Machine." *Literary Linguistic Computing*, vol. 18, no. 1, 2003, pp. 101–7.

–––, editor. *The Thomas MacGreevy Archive*. 2007. www.macgreevy.org/. Accessed 21 Aug. 2019.

–––. "Re-Envisioning Versioning A Scholar's Toolkit." *Digital Philology and Mediaeval Text,* edited by Ciula Arianna and Francesco Stella, 2007, pp. 93-102.

–––, et al. *Versioning Machine 5.0*. 2016. v-machine.org/. Accessed 21 Aug. 2019.

Vogeler, Georg. "The 'assertive edition'." *International Journal of Digital Humanities*, vol. 1, 2019, pp. 309–22. doi.org/10.1007/s42803-019-00025-5. Accessed 12 Oct. 2019.

Pierazzo, Elena. "Digital Documentary Editions and the Others." *The Annual of the Association for Documentary Editing*, vol. 35, 2014. scholarlyediting.org/2014/essays/essay.pierazzo.html. Accessed 21 Aug. 2019.

–––. "Facsimile and Document-Centric Editing." *Digital Editing of Medieval Texts: A Textbook*, edited by Marjorie Burghart. 2017. www.digitalmanuscripts.eu/digital-editing-of-medieval-texts-a-textbook/. Accessed 21 Aug. 2019.

What is Variance?

Towards a Model of (Variant) Readings

Elisa Nury

Abstract

In scholarly editing, more particularly in the context of collating various versions of a text, the definition of a variant reading is crucial. Yet, despite its importance, the meaning of a variant reading is often reduced to a "difference." The reason for such a vague definition is that what makes a variant can largely depend on the field of study: scholars of the Homeric oral tradition will consider different variants from scholars of medieval traditions or early printed texts, or from genetic critics. This contribution will focus on the modelling of a *reading*, arguing that formalizing this concept is necessary in order to define, and thus model, a *variant*. This article will also address digital representation of a reading by focusing on one implementation: the JSON data format used in conjunction with collation programs such as CollateX.

What is a version? In textual criticism, the term *version* may specifically describe a major rewriting of a work, possibly by the author. Here, however, we will consider versions in a

broader sense. The critical comparison–or collation–of different versions of one text is a necessary step during the preparation of a text-critical scholarly edition. Each version of the text is recorded in a document–or witness–and consists of readings, i.e., the particular word or words found at a given point in the text. In this context, a version is determined, amongst other characteristics, by the *differences* in the words found in the text, or *variant readings*. Variant readings are important since they provide valuable information regarding how versions are related to each other and how the text evolved through transmission. This article will focus on the modelling of *readings*, arguing that formalizing this concept is necessary to define, and model, variant readings. We will show how reading was a technical term that was used quite consistently through the ages, until it was defined with precision. Then we will establish the basis for a model by selecting important features of textual readings according to the previously examined definitions. These features, such as the textual content (or absence thereof), its size, and location in the text, will be discussed, raising various issues. This article will also address digital representation of a reading by focusing on one implementation: the JSON data format used in conjunction with collation programs such as CollateX. As we will see, the concept of variant readings may depend on the tradition of the text in consideration, and a variant in Homeric epic is different from a variant in a medieval tradition. The concept of variant is also dependent on the purpose of the comparison: a scholar attempting to reconstruct a stemma, or a linguist, may need to examine different variants. Therefore, a model of a reading should make it possible to distinguish different sets of variants depending on the context, and we will examine how the JSON implementation makes it possible with a few examples.

Let us consider the example of figure 1, where four versions of a sentence are aligned. When comparing the sentences of A, B C,and D, some readings can be considered equivalent in all four sentences, such as *The* or *upon*; other readings are different and change the meaning of the sentence: the absence of the adjective *bright* in sentence B, the triplet *star/sun/stars*, and the verbs with different tense (*shines* and *shone*). Finally, some readings are different, but may not alter the sense of the sentence (such as *worlde* and *world* or *sun* and *sunne*). Readings are thus divided between equivalent readings and different readings, and among the different readings a set of readings may be considered significant variant readings (see figure 2).

A	The	bright	star	shines	upon	the	world.
B	The		sun	shines	upon	the	world.
C	The	bright	stars	shone	upon	the	world.
D	The	bright	sunne	shines	upon	the	worlde.

Figure 1: Readings.

In the short collation extract of figure 1, there are four places where differences appear in the text. However, not all differences between readings are necessarily considered variant readings in any possible context. Scholarly opinions on this point range widely: from the view that every difference is a variant (Andrews) to considering only a limited number of "significant" differences to be variants, for instance, in the context of New Testament criticism, and therefore it is not enough to define a variant simply as a difference:

> The common or surface assumption is that any textual reading that differs in any way from another reading

in the same unit of text is a "textual variant", but this simplistic definition will not suffice. Actually, in NT textual criticism the term "textual variant" really means–and must mean–"*significant*" or "*meaningful* textual variant" (Epp 48).

In fact, the concept of *variance* has evolved with time and according to several theories. Since the nineteenth century, many scholars contributed to the development of a method for the establishment of genealogical relationships between manuscripts: the so-called Lachmann method. Maas in particular focused on a specific category of differences: shared errors, or indicative errors, can be used as a guide in order to assess the witnesses of the text and determine their relationships into a *stemma codicum*, or genealogical tree of textual witnesses.[1]

Greg separated variant readings into accidental and substantial, following the idea that some differences (substantials) have more importance than others (accidentals):

> [W]e need to draw a distinction between the significant, or as I shall call them, *substantive* readings of the text, those namely that affect the author's meaning or the essence of his expression, and others, such in general as spelling, punctuation, word-division, and the like, affecting mainly its formal presentation, which may be regarded as the accidents, or as I shall call them, *accidentals* of the text (Greg 21).

In the twenty-first century, scholars started to compare textual variants to DNA mutations and applied concepts from evolutionary biology and phylogenetics to textual criticism (Barbrook et al.; Salemans; Heikkilä). Lastly, in opposition to the distinction between accidental and

substantial variants, Andrews suggested a big data approach where every difference is a variant.

With the introduction of Lachmann's method, shared errors became the object of scholarly attention, and much work was done on the description and classification of the kind of errors committed by scribes who were copying manuscripts by hand. The cause of the error, as well as its conscious or unconscious character, is generally taken into account. Since the conscious modifications of scribal corrections were often attempts at improving or restoring the text, the terms *innovation* and *secondary reading* are frequently preferred to *error*. One of the most comprehensive review of errors was published by Havet, but other scholars have proposed other typologies of errors (Petti; Love; Reynolds and Wilson). These typologies often divide errors into four types: additions, omissions, substitutions and transpositions (Petti). When the scribe is consciously modifying the text, Petti (28–29) refers to scribal corrections as insertions, deletions and alterations instead of additions, omissions and substitutions. In parallel, many fields of study have offered their own definitions for variants according to their needs and their perspective on the text. From oral traditions such as Homeric epic to early printing, from medieval traditions to genetic criticism, from linguistics to phylogenetics, variants take many forms depending on the context: *multiformity* (Nagy), *early* or *late* states (Dane), variants at the sentence level (Cerquiglini), *open* variants, *type-2* variants (Salemans), and so on. The task of proposing a model for variant readings which would be suitable in any of the possible contexts, seems at best challenging, if not impossible. Rather than dealing directly with variants, this article will focus on modelling readings, especially textual readings. Not all readings are variant readings, but variants are always readings which differ in some respect from one another (see figure 2). Once readings have been modelled, variant readings could be

more easily modelled as a set of readings, with various criteria according to each discipline (V1, V2, V3). However, modelling those subsets will not be in the scope of this article. In order to propose a model for readings, we will first review the origins and usage of the term as well as its definitions in Section 1. The analysis of definitions will provide a first outline for a model, which will be discussed in Section 2.
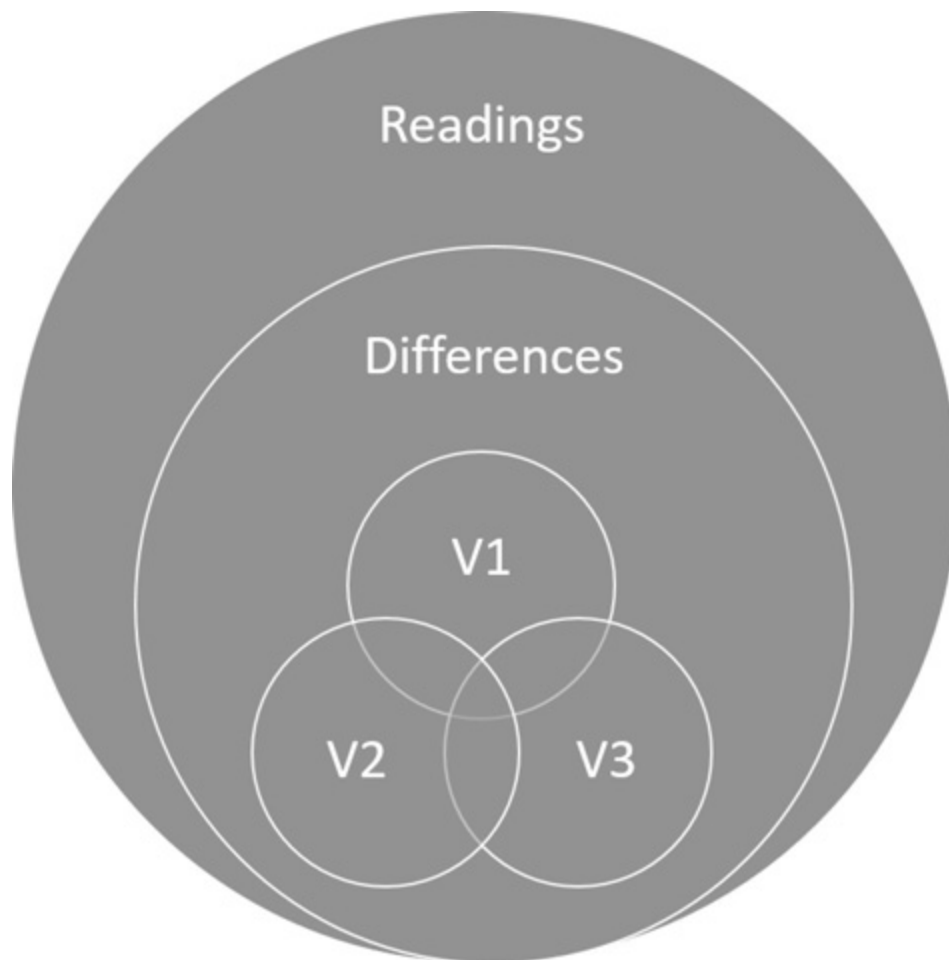

Figure 2: Readings, differences and variants.

1 Readings in context

Reading is a technical term that has long been used in the context of textual criticism and philology. It was already attested with Alexandrian critics: terminology included *graphe* (what is written), and *anagnosis* (what is read, a reading). The Latin equivalents are *scriptura* and the most common *lectio* (Montanari 26). The terms used by scholars of Antiquity imply a distinction between the words that are actually written on the page as opposed to the interpretation of the text. In English as well, a reading implies a form of interpretation; it could be read in more than one way. Here are a couple of examples where the words *scriptura* and *lectio* are used to qualify textual variation:

> *Obolus, id est, virgula iacens, adponitur in verbis vel sententiis superflue iteratis, sive in his locis, ubi* lectio *aliqua* falsitate *notata est, ut quasi sagitta iugulet supervacua atque falsa confodiat.* Isidore 1.21.3.

> The obelus, that is, a horizontal stroke, is placed next to words or sentences repeated unnecessarily, or by places where some **passage** is marked as **false**, so that like an arrow it slays the superfluous and pierces the false. (Barney et al.)

> *"Et idcirco inportunissime," inquit, "fecerunt, qui in plerisque Sallusti exemplaribus* scripturam *istam* sincerissimam *corruperunt."* Aulus Gellius 20.6.14.

> "And therefore," said he, "those have acted most arbitrarily who in many copies of Sallust have corrupted a **thoroughly sound reading**." (Rolfe)

Here the nouns *scriptura* and *lectio* have been emphasized, as well as the term which qualifies them. As these passages demonstrate, there was a strong focus in Antiquity on

whether a reading is corrupt or sound. When producing a new literary book, Hellenistic scholars used to correct a single copy of a work, instead of comparing as many copies as possible as modern editors do. This practice led Hellenistic scholars to become correctors of a specific work, and some experts compared them to editors (Montanari). Therefore, the need to distinguish between authentic and spurious readings arose, which may have motivated the dichotomy between sound versus corrupt readings, true versus false. The concept of variant reading, however, appeared much later during the Renaissance. In the Renaissance, Humanist scholars who were rediscovering and editing classical texts of Latin and Greek literature started to deploy technical terms that would become the base of the language of textual criticism. Silvia Rizzo's *Lessico Filologico degli Umanisti* provides invaluable information about the vocabulary in use amongst famous Humanists in the fourteenth and fifteenth centuries. By analysing their correspondence and publications, Rizzo was able to extract global definitions and explain what they meant when they used a given word. During the Renaissance, as Rizzo (209–13) shows, *lectio* and *scriptura* continued to be used as synonyms in much the same way as in Antiquity, for a passage of a text that can be read in a manuscript or an edition. Renaissance scholars would apply the term to readings from manuscripts as well as conjectures by other Humanists, and would mostly describe those readings as either correct (*recta, sincera*) or incorrect (*corrupta, mendosa*) according to their judgement. At the same time, the concept of *variant reading* started to be used more precisely with *varietas* (diversity) and in expressions where *lectio* or *scriptura* were used in connection with the adjective *varius*. Lorenzo Valla and Girolamo Avanzi have both used *varia lectio* and *varia scriptura* to describe a portion of text with different possible readings, as reported by Rizzo (213). Valla was accused by

Poggio of having presumptuously corrected a verse from Sallustius' first Elegy. Valla replied to Poggio that he did not emend Sallustius but merely chose one reading in a passage that varies (*varia scriptura*), even though the reading was attested only in very few manuscripts.[2] Another scholar, Avanzi, was asked for his opinion on a difficult passage from Catullus I, 9. He offers no solution of his own to emend the corrupted text, but he sends to his correspondent a list of conjectures (*varia lectio*) proposed by others.[3]

The usage of *lectio* and *scriptura* illustrates two contrasting approaches to readings and variant readings. Usually, a reading becomes a variant only when compared to another reading (Froger 80); variant also implies a deviation from a norm, one version of the text which may be chosen at random (Colwell and Tune 253).[4] On the other hand, a variant can be one among multiple possible alternatives, in a place where at least two witnesses disagree as to what the text is. Consequently, Colwell and Tune decided to refer not to variants, but to *variation-units*. This approach is shared by genetic criticism, which reject the existence of an invariant text, against which variant readings are compared (Biasi). In the twentieth century, formal definitions of reading can be found for instance in editing manuals, dictionaries or lexicons. Stussi defines a reading as "a passage from a transmitted text as it appears in a given witness"(Stussi 89).[5] A more precise definition of a reading is given by Froger, while describing one of the first examples of collation software:

> The form or content of the text in a given place is a *reading*, that is to say what we read at this location. Any manuscript, for instance the original, can be considered regarding its content as a collection or set of readings, which are the text elements at various

levels: chapter, paragraph, sentence, word, syllable, letter, and even punctuation or accents (Froger 9).[6]

This definition adds more precision: a reading is a textual element ('what is read'), and it can be of various scope, from the smallest punctuation marks to whole chapters. How can these definitions of a reading lead to a first example of a reading model?

2 Modelling a reading

The purpose of data modelling in the Humanities is to describe and structure information about real-world or digital objects in a formal way, so that this information becomes computable (Flanders and Jannidis 229–30) and so that it can be manipulated and queried with the help of a computer in order to answer questions. Ultimately, the purpose of modelling readings is to help determine if two given readings may be considered variant readings in a specific context. Flanders and Jannidis (234) suggest modelling textual variants in a scholarly edition by classifying variants according to some scheme, such as accidental versus substantial, or orthographical versus lexical, which corresponds to a consensus within the community.

   As we have seen, however, variants can represent something very different depending on the perspective (stemmatics, linguistics, etc.) and textual traditions (oral, medieval, early printing, and so on); therefore, readings need to be modelled independently of their function in textual criticism, but with enough information to decide what is a variant in those contexts. It may be helpful to consider the distinction between readings and variants in the framework of Sahle's wheel of text model (Sahle 45–49). Readings can be considered as a part of the text as

Document (TextD), whereas variants are part of the text as Version (TextF). The text as Version is further divided into subcategories, such as TextK, a canonical representation of the text which aims at identifying the best (true) text. With this framework in mind, the characterization of readings as authentic or corrupt does not make a good model for readings, since it represents rather variants than readings. Therefore, the more recent definitions of readings may provide a better starting point to the model than the true/false distinction previously applied to readings. Models are simplified representations of an object of study, a selection of features among all available (Pierazzo 44–45). From the overview of the term reading provided in the previous section, in particular the definition of Froger and Stussi, features which apply to a reading can be inferred, namely that a reading:

- conveys textual content;
- has a precise location in the text (also referred to as *locus*);
- can occur at any level of the text, and thus have various sizes;
- is transmitted by a witness.

2.1 Issues

These features need to be discussed in more detail. For instance, is it too restrictive to limit a reading to textual content? What about decorations, mathematical diagrams and other non-textual elements? Historians of Greek, Arabic or Egyptian mathematics have acknowledged the need to collate and critically edit mathematical diagrams instead of simply providing corrected figures to fit modern standards. Raynaud created a stemma for the *Epistle on the Shape of the Eclipse* by Ibn al-Haytham, a mathematical treatise from the eleventh century, using the mathematical diagrams