**BHATTACHARYYA —** Approaching

<0 figures>

**Approaching Textuality with the Metaphor of the Digitized Workset**

Bhattacharyya, S. and Downie, J. S.

While much of digital humanities considers the affordances provided by computational tools, a less noted and more serendipitous consequence of its emergence is the availability of new metaphors for rethinking perennial philosophical questions. The concepts underlying the *workset*, a notion that has emerged out of the Workset Creation for Scholarly Analysis project of the HathiTrust Research Center, provide such a metaphor.

The *workset* is a collection, typically scholar-built, consisting of a select subset of digitized texts from a larger digital library’s collection (Fenlon et al., 2014). Important for our argument in this paper is the property of worksets that they need to be both stable and dynamic simultaneously. On the one hand, the workset must be coherently stable for the purpose of citability; if the researcher has performed computational analysis on the workset and published the result thereof, then the integrity of the content of the workset at the moment of that analysis needs to be maintained citably for the purpose of reproducibility of results. On the other hand, since the workset itself is a part of a larger digital library, individual elements of the workset are, conceptually speaking, pointers to the corresponding elements held in the library’s collection; those pointed-to elements in the library are mutable or dynamic—for example, the scanned-and-OCR-ed copy of a pointed-to volume in the larger digital library may, at a later point in time, be replaced by a different OCR-ed copy because the latter is of better quality, and a pointer to a volume may end up pointing to null because the library may have had to remove access to the pointed-to volume in response to a takedown notice from a publisher. We argue that this simultaneous, and antagonistic, demand on a workset, of *citability as well as mutability*, is a useful (though imperfect) metaphor for the contradictory qualities that text itself embodies. Additionally, we argue that the analogy between the notion of a workset and the notion of text itself, by providing a common unifying framework with which to think about such activities as editorial scholarship and anthology building, can help problematize the distinction between seemingly disparate aspects of textuality.

Bernard Cerquiglini has pointed out that not only did the modern idea of a text as standard and definitive, in which ‘every written word’ is a reflection of a supposedly immutable originary authorial intention, *not* have a strong counterpart in antiquity (1999, p. 4), but also that, at least for the first two centuries after the first written works in French started to appear, there was, in mediaeval France, an especially heightened incidence of variance in the written text: ‘in the Middle Ages, literary work was a variable’ (1999, 32); according to Cerquiglini, this suggests that, especially in any nascent literature, variance *is* the natural condition of existence of written texts (1999, 21). Elements of the text were mutable, though mutability of text declined over time with progressively wider dissemination of the technologies of textual production and, especially, of printing. Members of the Prague School (Bogatyrëv and Jakobson, 1982, 42) and Russian formalists such as Vladmir Propp (Propp, 1958, 1) have pointed out that this quality of variance has been retained in the genre of the folktale and its retellings. At the same time, as Propp also demonstrated in the case of Russian, functions in folktales serve ‘as stable, constant elements’, their number is ‘limited’, and their sequence within folktales often tend to be ‘identical’ (20). Thus, a folktale, in spite of the variance that its constituent elements might exhibit, nevertheless has a typically stable and easily recognizable structure. A folktale, arguably a seminal embodiment of the textual condition, is thus—if one were to employ the terminology associated with a workset—both a *mutable* and a *citable* object. The analogy breaks down in one respect: a folktale’s structure is *narrative*, that is, a *relational* order obtains between the constituent elements of its text, whereas a workset (in its essential implementation) does not necessarily have any notion of internal ordering—one of the reasons why the metaphor is not perfect.

We argue that this interesting, though imperfect, analogy between worksets and text may be useful, by foregrounding the notion of mutability, in approaching the long-standing debate among literary theorists and computation/information-oriented philosophers regarding notions of textuality. For example, the view from computation/information-oriented philosophy, influenced by technologies such as TEI, is that ‘text is an ordered hierarchy of content objects’—the so-called OHCO thesis (Renear, 1997, 118); a typical view of literary theorists such as Jerome McGann, on the other hand, is that texts, especially poetic texts, are ‘recursive structures built out of complex networks of repetition and variation’, and hence cannot be accommodated within the OHCO paradigm (Hockey et al., 1999). David Golumbia diagnoses this disagreement in perspective as irreconcilable (2009, 107), while Patrick Sahle notes how the onto-epistemological peculiarities of the OHCO paradigm help the articulation of some among various pluralistic textual concepts more than others (2013, 396)—versioning being among the latter (Schreibman, 2002). We argue that the emphasis on mutability imposed by thinking of text in terms of the metaphor of the workset is an instance of how a computationally inspired framework can accommodate variation. The variation and mutability in the workset paradigm is diachronic, whereas the variation that McGann refers to is synchronic; we argue that any synchronic variation is representationally equivalent to a set of diachronic variations.

The workset metaphor provides a way to conceptualize the notions of the two different kinds of editorial scholarship—editing of a single text, and the creation of an edited anthology of multiple texts—within the terms of a common framework. In a 1995 essay, Gunter Martens argues that ‘text’, in the editorial sense, should be understood and interpreted as a process unfolding in time, ‘so that the “text” is understood and interpreted as an historical process’ (1995, p. 195). Martens’ subject is the writer sitting in front of a piece of paper, writing, rewriting, and changing the text. Editorial scholarship, Martens argues, must produce neither the beginning point nor the endpoint of that process, but the process itself as it unfolds over time, that is, as the text mutates. Clearly, such an editorial practice is impossible to create, except as a utopian ideal, within the paradigm of printed editions, but it can be accomplished relatively easily using the conceptual idea of worksets. The historical process that the text embodies can be conceptually approximated by the same underlying mechanisms in a workset that provide support for its mutability. In addition, if we consider this idea of text as process also in its *social* dimension, then different versions of a text (as, for example, a book going through various editions, with its textual content undergoing corresponding modifications as the social context of the book changes), constitute mutations over time, and can be accommodated within the conceptual notion of the workset.

We can stretch this idea further. Different translations of a book, or even a sequence of books in which one book shows a strong influence of another, can arguably be considered as constituting a chronological process of mutation of the same ‘text’—a text mutating through different versions of itself. If we stretch the notion of ‘version’even further, then at the very limit of the notion of ‘version’, one could arguably consider books sharing a common theme as different ‘versions’ of the *same* book. The study by Fenlon et al. (2014) of how scholars organize their worksets reveals that scholars often have their workset constituted around a common theme. Such a workset built around thematic commonality is akin to an anthology of different texts, rather than to different versions of the same text. However, we will show that the difference between different versions of the same text, and an anthology of thematically similar texts, is a slippery difference. What conventionally distinguishes a collection of texts as versions of each other from an anthology of similarly themed material thrown together is a corresponding distinction between the implicit notions of temporally mediated causality and spatial synchrony. We will show how the ‘isGatheredInto’ property, proposed by Wickett et al. (2014) to mediate between part-of and member-of types of relationships (and to subsume them both), can provide a way to conceptually unify these twin, but seemingly disparate, ideas of versions-maintenance and thematized collectability (operationizable as the editorial/textual-critical practice of variorum creation, and the curatorial practice of anthology making, respectively). The first of these ideas/practices concerns temporality, and the second concerns spatiality. Such a unification is an illustration of how the consideration of systems that, from an engineering point of view, satisfy a requirement (in this case, the requirement of meeting the simultaneous and oppositional demands of citability and mutability) can *also* be helpful, by means of the application of the same principles, in thinking about the metaphysics of textuality, within the context of which literary theorists have noted a similar unification of temporal ‘play’ and spatial ‘structures’ (McGann, 2001, 178).

**References**

**Bogatyrëv, P. and Jakobson, R**. (1982). Folklore as a Special Form of Creativity. In Steiner, P. (ed.), *The Prague School: Selected Writings, 1929–1946.* Austin: University of Texas Press, pp. 32–46. (Originally published as Bogatyrëv, P. and Jakobson, R. [1929]. Die Folklore als eine besondere Form des Scaffens. *Donum natalicum Schrijnen*. Nuijmege, pp. 900–913.)

**Cerquiglini, B.** (1999). *In Praise of the Variant: A Critical History of Philology.* Betsy Wing (trans.). Johns Hopkins University Press, Baltimore. (Originally published as *Éloge de la variante: Histoire critique de la philologie*. Paris: Éditions du Seuil, 1989.)

**Fenlon, K., Senseney, M., Green, H., Bhattacharyya, S., Willis C, and Downie, J. S.** (2014). Scholar-Built Collections: A Study of User Requirements for Research in Large-Scale Digital libraries. *Association for Information Science & Technology (ASIS&T)*, annual meeting, Seattle, WA, 31 October–4 November 2014.

**Golumbia, D.** (2009). *The Cultural Logic of Computation*. Harvard University Press, Cambridge, MA.

**Hockey, S., Renear, A. and McGann, J.** (1999). What Is Text? A Debate on the Philosophical and Epistemological Nature of Text in the Light of Humanities Computing Research. In *Proceedings of the 1999 Joint Annual Conference of the Association for Computers and the Humanities and the Association for Literary and Linguistic Computing (ACH-ALLC ’99)*, Charlottesville, VA, 9–13 June 1999, http://www2.iath.virginia.edu/ach-allc.99/proceedings/hockey-renear2.html.

**Martens, G.** (1995). What Is a Text? Attempts at Defining a Central Concept in Editorial Theory. In Bornstein, G., Pierce, G. B. and Gabler, H. W. (eds), *Contemporary German Editorial Theory.* Ann Arbor: University of Michigan Press, pp. 209–31.

**McGann, J.** (2001). *Radiant Textuality: Literature after the World Wide Web*. Palgrave, New York.

**Propp, V.** (1958). *Morphology of the Folktale*. Scott, L. (trans.), Pirkova-Jakobson, S. (ed.). Bloomington: Indiana University Research Center in Anthropology, Folklore, and Linguistics. (Originally published in Russian as *МОРФОЛОГИЯ СКАЗКИ*, 1928.)

**Renear, A.** (1997). Out of Praxis: Three (Meta)theories of Textuality. In Sutherland, K. (ed.), *Electronic Text: Investigations in Method and Theory.* Oxford University Press, Oxford.

**Sahle, P.** (2013). *Digitale Editionsformen. Zum Umgang mit der Überlieferung unter den Bedingungen des Medienwandels. Teil 3: Textbegriffe und Recodierung. [Preprint-Fassung].* PhD thesis, Universität zu Köln, http://kups.ub.uni-koeln.de/5013/1/DigEditionen\_3\_online.pdf.

**Schreibman, S.** (2002). Computer-Mediated Texts and Textuality: Theory and Practice. *Computers and the Humanities,* **36**(3): 283–93.

**Wickett, K. M., Doerr, M., Meghini, C., Isaac, A., Fenlon, K. and Palmer, C.** 2014. Representing Cultural Collections in Digital Aggregation and Exchange Environments. *D-Lib Magazine,* **20**(5/6) (May/June), http://www.dlib.org.dlib.may14/wickett/05wickett.html.