J.R. Mattison, University of Georgia, and Hannah K. Ryley, University of Oxford

Introduction: Big Data and Manuscript Studies: Some Reflections on Methods

Abstract: This introduction considers the approaches to, reasons for, and impacts of using "big data" methods to study medieval manuscripts. This introduction to "big data" and manuscript studies offers a brief history of the field, and its terms, before laying out some of the benefits, challenges, and drawbacks of large-scale work. Looking at lots of books presents an intellectual clash: the desire to know more about what we are looking at and the desire to see *everything* often runs up against the knowledge (and the fear) that we cannot know everything. We argue that undertaking large-scale work necessitates reflecting not only on reasons for counting, analyzing, and studying manuscripts at scale, but also on how we count and why. Methods always matter, and reflecting critically on methods matters even more when there are still foundational questions to be answered about medieval manuscripts.

Medieval bibliophiles desired books. Lots of them. The thirteenth-century author of the *Philobiblon* Richard de Bury was "reported to burn with desire for books, and especially old ones" and details his ingenious schemes for acquiring as many of them as possible.[1] Bibliophiles like de Bury typically sought out old books for the ancient and biblical knowledge they contained. They were nostalgic for a lost past and wanted to study the 'truths' held in old books. Hoarding more and more precious books shored up that true knowledge safely, or so their thinking went. But–often in the same breath–bibliophiles also bemoaned the loss of old books, their sorry physical state, and the difficulties of accessing them, especially in any great quantity.

Early bibliophiles envisioned within their pursuit of lots of books a pursuit of knowledge, a pursuit of an inaccessible completeness. Whether and how to read such books seems a less important question to de Bury than simply how to amass them. And this relationship between inaccessible knowledge and lots of books is only amplified when more and more books are brought together. The scale of the unknown comes into sharper view. Surely studying more books offers greater access to a lost past, and to find the truths that it holds? Yet no matter how many books are gathered, the past is never perfectly recoverable. Building ever larger collections of older books cannot undo the long history of manuscript loss and destruction. Like medieval bibliophiles before them, modern codicologists desire access to (if not perhaps acquisition of) lots of old books. Then as now, scholars of manuscripts run up against similar challenges when confronted by the textual and material complexities of big collections of books, not least the keen awareness that loss is insurmountable and that the manuscript record will always be incomplete.

Quixotically, then, an understanding of loss–both intellectual and physical–is always implicit in manuscript abundance. John Aubrey, the seventeenth-century gentleman scholar and antiquary, grieved to see manuscript leaves being used to bung up barrels of ale, cover schoolbooks, wrap gloves, and scour guns. He bemoans that: “In my grandfathers dayes, the Manuscripts flew about like Butterflies: All Musick bookes, Account bookes, Copie bookes &c. Were covered with old Manuscripts [...] Before the late warres a World of rare Manuscripts perished here about.”[2] Here, manuscripts are abundant, but only in their fragmentariness. Alongside such accounts of loss, sometimes even large collections of books could not satisfy the exacting desires of bibliophiles. Writing from London in 1421, Poggio Bracciolini lamented that libraries in Europe held many books, but not the humanist works he sought: “On this trip of mine I visited a monastery that is older than all the others in England and more magnificent. I inspected the library thoroughly: there are many books but none for us. I saw other libraries besides in many places, all full of nonsense."[3] Although he gathered materials from many books and compared their texts, Bracciolini knew that he would never be able to find everything he hoped to know. Sometimes a lot feels like it is not enough.

Looking at lots of books therefore presents an intellectual clash: the desire to know more about what we are looking at and the desire to see *everything* often runs up against the knowledge (and the fear) that we cannot know everything. While modern manuscript scholars have more varied interests than de Bury, Aubrey, and Bracciolini, they nonetheless encounter many of the same frustrations. The desire to know about books in quantity yields further questions: how many manuscripts are there in total? How many have been lost? How many survive in fragmented states? How do we study texts in their manuscript contexts? Looking at, counting, quantifying, and classifying lots of books certainly represents a source of knowledge and helps answer some of these questions. However, such processes also raise the issue of *how* we do so. We should be asking not just "how many," but also "how do we count" and "why do we count."

The essays in this collection reflect on the methodologies of analyzing larger groups of books. Such considerations are particularly important to the field of English manuscript studies, which has often been hesitant to deploy large scale approaches and relied on case studies.[4] The essays collected here, rather than offering definitive how-to's, ruminate on the challenges, questions, and opportunities that arise when studying lots of books. Our contributors, working predominantly in the field of Anglophone manuscripts, discuss their own processes of exploring and gathering archival, physical, and digital materials that have often not necessarily been studied in large scale ways. They provide a cross-section of perspectives that consider issues of terminology, technology, and disciplinary silos. These methodological case studies model ways of dealing with the increasing amounts of data about books available on- and off-line. Contributors' corpora range in scale from manuscripts of Chaucer's *Troilus and Criseyde* to 27,000 fragments of books from medieval Scandinavia. While a quantitative approach figures in some of these essays, it is not the only method for studying many manuscripts. Through short, focused explorations, the essays gathered here offer insight into how the field has developed, and where it might progress.

This special issue, focused on English manuscript studies, builds on decades of studies into large groups of manuscripts. In manuscript catalogs and in editions of popular medieval texts, in studies of medieval libraries and in catalogs of dated and datable manuscripts, lots of books are grouped together as veins of evidence. Paleographers, codicologists, textual scholars, and other book historians have recovered knowledge about the production, dissemination, and use of manuscripts. The importance of large-scale approaches seems self-evident: it is only by examining and synthesizing hundreds of manuscripts that we can develop a picture of how–for instance–scripts changed over time and place, how book binding evolved, or how medieval people read. The study of manuscripts in quantity is arguably tied to the origins of the historical study of manuscripts in the seventeenth century. Jean Mabillon, the Benedictine monk who was the first to publish a study of paleography, brought together many examples of scripts from libraries across Europe.[5] In more recent decades, quantitative approaches have been particularly associated with European–as opposed to British and American–methodologies. Studies by J. P. Gumbert, Ezio Ornato, Carla Bozzolo, and Albert Derolez contribute to our foundational understanding of medieval European manuscripts.[6] Large scale work, as well as (the often related) quantitative approaches not only lay out the motivations for studying manuscripts quantitatively, but also the possibilities of new discoveries. For instance, Gumbert, with help from unnamed assistants, synthesized 2,395 manuscripts to determine how page sizes in manuscripts changed over the thirteenth, fourteenth, and fifteenth centuries.[7] J.A. Szirmai examined "about one thousand" bindings in order to describe and define binding styles across Europe.[8] Sonia Scott-Fleming drew on 86 manuscripts to distinguish the characteristics of French and English penwork flourishing.[9] Kathleen Scott selected 50 manuscripts from among some 200 to establish a timeline of features in English manuscript borders.[10] Erik Kwakkel used 342 manuscripts to chart the transition from Caroline miniscule to Gothic script.[11] Devani Singh analyzed 83 manuscripts and 38 incunables of Chaucer to elucidate the poet's early modern reception.[12] There are far more studies than can be listed here; this sampling shows just how varied contributions that use large groups of books can be. These studies deploy a variety of methods, from the analysis of graphical models of data to the more qualitative analysis of trends derived from a corpus. In drawing on assemblies of books, selected for time, provenance, and contents, these authors and many others provide a sense of the shape of the medieval book.

Additionally, the approaches described in this special issue are indebted to recent work in the digital humanities. Digital tools and techniques that draw on the data of hundreds or thousands of manuscripts offer new ways of organizing, analyzing, and sharing such work.[13] The digitization of manuscripts and library catalogues have been particularly important. As Aditi Nafde and Emma Gorst have argued, digital manuscripts "[enable] scholars to read more, to read faster, and to read differently."[14] Certain forms of study that once required weeks, months, or years can be completed in hours, even as, Bridget Whearty and Dot Porter remind us in their contribution, digital manuscripts are not simple replacements for physical books. Repositories of images, from the Vatican Library DigiVatLib (which stands at 28,312 digitized manuscripts at the time of writing but aims for 80,000) to Flickr albums of smaller institutions, represent forays into the world of "big data" and manuscript studies.[15] In turn, these institutions that house books produce ever more new data about them, through new catalogs, bibliographies, metadata, and more.

Yet, the connection between medieval studies and digital humanities work, again longstanding, raises important ethical questions about the way we approach lots of books: Roberto Busa's index to the works of Thomas Aquinas, begun in 1949 and made up of over 10 million indexed terms, is often called the first digital humanities project. The index is, ultimately, based on manuscript data turned into a different form. But as Dorothy Kim and Arun Jacob demonstrate, Busa's index, with its connections to Mussolini's Italy and military technology, creates an origin story for digital humanities rooted in white supremacy and fascism.[16] These and other critical approaches to the digital humanities prompt necessary reflections on how big data and digital humanities intersect with forms of bias, structures of power, racism, sexism, and nationalism.[17] The large amounts of data about manuscripts made possible through digital humanities projects are connected to these histories. Acknowledging big data and manuscript studies' links to digital humanities reminds us not only to think critically about our works' attachments to systems of oppression, but also the sources of our study, the limits of our knowledge, and the kinds of questions we ask the data to answer. In addition to "how" and "why" do we count, we might ask the related questions, "what *can* we count and why"?

This overview demonstrates how looking at many manuscripts touches on a variety of fields, practitioners, and ways of knowing. Acknowledging these histories, we re-situate newer developments and methodologies. Rather than sharing the results of their studies of many books, our contributors focus instead on how they went about their studies. Methods always matter, and reflecting critically on methods matters even more when there are still foundational questions to be answered about medieval manuscripts. Indeed, as some of our essayists point out, common assumptions about manuscripts may yet be overturned by quantitative studies.

While quantitative approaches are used in manuscript studies across several fields and disciplines, many of the contributors here attend to English medieval manuscript studies in particular. The production of books in England in the Middle Ages differed from production elsewhere, from choices of script to the organization of scribal labor, and the study of such manuscript has also evolved differently. As Daniel Sawyer has commented and Michael Johnston elaborates in his essay here, English manuscript studies has yet to embrace the quantitative approaches that are more familiar to French or Dutch or Italian manuscript studies.[18] Because this field and its very boundaries are still developing, particularly for Middle English studies, there is still much to do. Now is a moment for reflection. We acknowledge that this special issue's partial focus on English manuscripts and on England represents a clear bias. However, this lens presents an opportunity to reflect critically on a well-established field of study, its quantitative methods, its achievements, and its futures. In this special issue, we ask not only *what* we do when we look at a lot of books, but also *who* looks, *how* we look, and *why* we look.

We employ the term "big data," but what do we mean by that term? The same question has been considered in other fields, like sociology.[19] For the humanist, "big" data may mean tens or hundreds or even thousands of items. The question takes on a slightly different meaning when applied to the unique, hand-written object. By the standards of other fields, the quantities of manuscripts that we have do not rise to the level of "big data." Even though 27,000 fragments might seem like a big number, in comparison, studies of printed books can easily incorporate tens or hundreds of thousands of items.[20] HathiTrust, one of the most used digital repositories of printed material, contains some 17 million titles. The kind of big data used in the sciences and social sciences group millions of data points.[21] Studies of seventeenth-, eighteenth-, nineteenth-, and twentieth-century print culture can mine vast quantities of books–digital and physical–that simply do not exist for the medievalist interested in manuscripts. By comparison, a study of a few hundred manuscripts seems paltry, even when such a study may be the work of years, close analysis, and hours in the library or in front of the computer screen. Instead, uniting big data and manuscript studies means rethinking our sense of scale. Even a single manuscript can contain within itself large data sets: of scribal spellings or paleographical features, forms of corrections and numbers of folios. ‘Bigness’ and ‘smallness’ can be produced through the different scales of the material object of the book itself, from the formation of individual letters to the collection of many volumes.

Like “big,” “data” similarly conjures a range of associations, such as mathematical precision and verifiable truths. We might think that data should, at least, be countable. Yet the size of a data set based on medieval manuscripts will be limited by the known, the cataloged, and the discoverable. Further, counting, and counting manuscripts in particular, can be a more fraught exercise than it seems. How, for instance, would you count the number of surviving John Lydgate’s *Fall of Princes* manuscripts? Do manuscripts with only some of *Fall* count as much as those with most of the poem? What about fragments? Lost manuscripts? Turning manuscripts into data is rarely as simple as numbering known books in succession. Once gathered, any collection of information needs organizing and analyzing. Drawing on Yin Liu’s work on “medieval codes,” we can even approach medieval texts in manuscript contexts as a form of data that can be processed and transmitted in ways other than human reading.[22] But, as Gumbert noted, looking at many books might be “a foreign country” to scholars ill at ease with numbers, formulas, and spreadsheets.[23] We often lack training in the kinds of statistical and digital approaches to data that are necessary for analysis. Turning manuscripts and their texts into usable data seemingly robs them of their “defiantly individual impulses,” the idea that each manuscript is a unique object.[24] We might therefore be inclined to contrast attention to single manuscripts or exceptional scribes with the study of tens, hundreds, or thousands of books. Yet, Gumbert lays out further, “Knowledge about [many books] will in its turn help in understanding single books.”[25] At the same time, even knowing what kinds of questions to ask of our data are informed by qualitative work. Individual manuscripts and large groups of manuscripts feed into each other, each helping with the study of the other. The deep study of individual manuscripts and qualitative approaches are not necessarily opposed to the broad study of manuscripts as data.

Working with manuscript data often means making decisions about the lost, fragmentary, incomplete, and ideal, decisions that may not necessarily be relevant to other fields. The suitability, meaning, and limits of terms like “big data” and its use in the field of “quantitative codicology” is worth investigating. In the search for complete knowledge, some things are unknowable for now and some things will always remain unknown. Because manuscript studies operates differently from other fields, our “big data” looks different and requires new approaches.

Closely related to questions of scale and countability is the question of lost books. Medieval England might have lost more of its vernacular medieval manuscripts than other locations in Europe.[26] From medieval reuse to the destruction of the Reformation, over-enthusiastic antiquarians who cut up books to World War II bombings, medieval books have suffered. As de Bury knew, any group of manuscripts we can now locate will differ from what was once available. Indeed, of the five cartloads of books sold by de Bury's executors, no book belonging to him is known to survive.[27] Loss and the threat of loss are an integral part of manuscript culture. But rather than lamenting these losses, as Bracciolini and Aubrey did, big data approaches offer ways of examining it critically.[28] Through the application of statistical models, we can gain a sense of the scale of loss, or extract dominant trends from what remains. Big data study does not undo loss, but it provides a set of strategies for thinking with loss.

The methodological practices of big data study reveal, too, that loss is not a purely medieval or early modern phenomenon, but one that affects the knowledge that we have about manuscripts and the availability of that knowledge. Many collections may not be accessible, or fully cataloged, or cataloged to modern standards, so that even answering seemingly simple questions like, "Which texts are in this manuscript?" or "what kind of parchment or paper was this manuscript written on?" may not be possible. Even as it seems that more manuscripts are increasingly accessible in digital formats and that it has never been easier to look at and compare books from a distance, we should always be mindful that such access is neither shared equally nor permanent. Which manuscripts seem to merit digitization and how are contingent on any number of factors.[29] Describing the practice of bibliography more generally, Derrick Spires reminds us that "lists, catalogs, organizational principles, bookshop displays, and the like are all arguments about what matters, how they matter, and ways of knowing."[30] Manuscript studies navigates such aporia in various ways, formulating approaches that are both similar to and different from the study of lots of books in other time periods. Ultimately, the practices of manuscript study are distinct, and scholars of medieval manuscripts cannot import wholesale the methods of other disciplines. Here we focus on the relationship of manuscript studies to questions of scale, a subset of ways of studying old books. These methods can and should be used alongside other qualitative and theoretical ones, including emergent feminist, queer, and trans approaches to bibliography and codicology.[31]

Moreover, the application of big data methods to manuscript studies should ask: what are the ethical dimensions of the quantitative methods currently in play? On a practical level, physically examining many manuscripts is an expensive and time-intensive process with ecological ramifications. There are real-life challenges for academics wishing to undertake such work, which might involve extensive travel and even relocation, putting pressure on caring and teaching responsibilities. Visa requirements can make access to libraries prohibitive for certain researchers, especially those from the Global South wanting to travel to the United States and the United Kingdom. Furthermore, jet-setting across the globe in pursuit of old books in a warming and ecologically vulnerable world is increasingly irresponsible. Now is the time to consider other ways of examining lots of books that draw on collaborative, digital, and ecologically-minded methods. We do not pretend to have a perfect solution, but we can think of alternatives like working less individually and more collectively to obtain relevant information that would reduce travel. Although humanist academic training elevates the scholarly work of the individual, collaborative and communal approaches borrowed from other disciplines may hold the key to envisioning future methodologies. We worry, too, that as funding sources for graduate students, non-tenure track, contingent, and independent scholars disappear, the kinds of scholars who are able to undertake big data approaches to manuscript studies will be yet more limited. Big data approaches to manuscript studies expose issues of access and privilege at every point in the development of a prospective project, from acquiring the requisite training in codicological skills to plan such a project to the ability to see lots of books at all. This narrowing of voices and perspectives and experiences is detrimental to a field that is still growing. The use of digital technologies, and the issues of labor, sustainability, dissemination, and access should all prompt us to pause and think again about how we study manuscripts. While quantitative approaches may well better represent more quotidian–or even boring–aspects of manuscripts, counting things is not always democratic in nature. After all, the methods we choose shape our interpretations and can re-entrench existing prejudices and biases. Big data and manuscript studies is entangled with the value and future of the humanities as a whole.

In the nine essays that follow, our contributors tackle these and other questions, offering their reflections, provocations, and manifestos. These essays range from broad assessments of the state of the field to focused case studies. Michael Johnston addresses the need for large scale approaches to Middle English manuscripts studies by identifying three lacunae in our knowledge. By working towards filling these gaps, English manuscript studies can align better with the work of Continental scholars and increase our knowledge of England's manuscripts in other languages. Daniel Sawyer turns to the problem of loss, reflecting on the issues and possibilities posed by the masses of lost, early books. Tackling manuscript studies in the face of loss is an area particularly suited to large scale work. Julia King picks up on some of these themes of loss, but also offers a number of cautions. Data, King argue, can be misleading in its apparent inherent empiricism. Confronting this 'quantitative bias' means equipping ourselves with statistical knowledge and continuing to rely on humanistic qualitative contextualization. Daniel Wakelin similarly thinks about the relationship between big data methodologies and historicized ways of understanding manuscripts. While macro approaches to manuscripts can yield new knowledge, we risk emphasizing the variant over the ordinary, the strange over the familiar. This essay also makes the case for recuperating the less exceptional by reminding ourselves about the difference between medieval and modern knowledge about manuscripts. In

“How to Look at Lots of Books of Hours,” Jessica Brantley examines how our knowledge about the so-called ‘medieval bestseller’ changes when we examine books of hours at different scales. Brantley lay out the possibilities offered by collaborative work. Aditi Nafde shifts our perspective to the tools that modern scholars use to study lots of books in “Looking Across Categories.” She demonstrates that while manuscript and early print studies have much to say to one another, the tools used by scholars in the two fields–if they really are separate fields–are often incompatible. Jaakko Tahkokallio and Ruen-chuan Ma and provide reflections and case studies. Ma studies the idea of manuscript data, balancing the study of multiple books against the reading experience offered by individual manuscripts. Focusing on a single stanza from *Troilus and Criseyde*, Ma examines different ways of manipulating the manuscript page into forms of comparable data. Tahkokallio explores how large numbers of fragments have the methodological potential to tell us more about medieval manuscripts and book production. Fragments, in contrast to whole books, represent a remarkable source of potential knowledge. Finally, Dot Porter and Bridget Whearty highlight the sometimes fraught relationship between digital manuscripts and big data and manuscript studies. Reflecting on current digital tools and future possibilities seemingly offered by artificial intelligence, they provide a number of timely cautions and best practices to keep in mind for those wishing to work with digital books.

****

The essays that make up this special issue pose focused explorations into big data and manuscript studies. Several of these essays were first read at the 2022 New Chaucer Society Biennial Congress. We are grateful to our original panel and our audience at that event for their enthusiasm and insight. Ranging from considerations of disciplinary boundaries to the proper handling of metadata, our contributors both undertake new approaches and provide timely reflections.

NOTES

[1] De Bury, *The Philobiblon*, 128.
[2] Aubrey, *Memoires of Naturall Remarques*, 221.
[3] Bracciolini, "Letter XV," 55.
[4] Johnston and Van Dussen, *Medieval Manuscript Book*, 3.
[5] Grafton, *Inky Fingers*.
[6] Gumbert, "Fifty Years of Codicology," 505–26; Ornato, *La face cachée*; Bozzolo and Ornato, *Pour une histoire du livre manuscrit*; Derolez, *Codicologie des manuscrits en écriture humanistique*.
[7] Gumbert, "The sizes of manuscripts."
[8] Szirmai, *Bookbinding*, xi.
[9] Scott-Fleming, *Pen Flourishing*, 11.
[10] Scott, *Dated and Datable*.
[11] Kwakkel, "Biting, Kissing and the Treatment of Feet."
[12] Singh, *Chaucer's Early Modern Readers*.
[13] E.g., Besamusca et al., *Multilingual Dynamics of Medieval Flanders*; Warren and Weijer, *Reimagining History*.
[14] Aditi Nafde and Emma Gorst, "Digital Medievalist," 148–149.
[15] *DigiVatLib*.
[16] Kim, "Embodying the Database," 170–87; Jacob, "Punching Holes."
[17] Kim, "Introduction."
[18] Sawyer, *Reading English Verse in Manuscript*, 10–11; Da Rold and Maniaci, "Medieval Manuscript Studies," 3.
[19] Weller et al., *Big Qual*, 4–8.
[20] See, for instance, Underwood, *Distant Horizons*; Bode, *A World of Fiction*; Bode, *Reading by Numbers.*
[21] Pietsch, *Big Data*.
[22] Liu, "Ways of Reading," 9–10.
[23] Gumbert, "Fifty Years of Codicology," 525.
[24] Hanna, "Miscellaneity and Vernacularity," 37.
[25] Gumbert, "Fifty Years of Codicology," 508.
[26] Kestemont et al., "Forgotten Books."
[27] Gillespie, "Books and Booklessness in Chaucer's England," 85.
[28] For attempts to conjecture how much has been lost, see Kestemont et al., "Forgotten Books."
[29] Treharne, *Phenomenal Book*, 195–212.
[30] Spires, "Liberation Bibliography," 19.
[31] Ozment, "Feminist Bibliography"; Magnani and Watt, "Queer Philology"; Sargan, "Trans Codicology."

WORKS CITED

Aubrey, John. *Memoires of Naturall Remarques in the County of Wilts*, Royal Society MS 92.

Besamusca, Bart, et al. *Multilingual Dynamics of Medieval Flanders*, University of Utrecht, https://multilingualdynamics.sites.uu.nl/project-team/.

Bode, Katherine. *A World of Fiction: Digital Collections and the Future of Literary History*. Ann Arbor, MI: University of Michigan Press, 2019.

Bode, Katherine. *Reading by Numbers: Recalibrating the Literary Field*. London: Anthem Press, 2012.

Bozzolo, Carla, and Ezio Ornato. *Pour une histoire du livre manuscrit au moyen âge: trois essais de codicologie quantitative*. Paris: Editions du Centre national de la recherche scientifique, 1980.

Bracciolini, Poggio. "Letter XV." In *Two Renaissance Book Hunters: The Letters of Poggius Bracciolini to Nicolaus de* Niccolis, edited and translated by Phyllis Walter Goodhart Gordon. New York: Columbia University Press, 1991.

Da Rold, Orietta, and Marilena Maniaci. "Medieval Manuscript Studies: A European Perspective." In *Writing Europe, 500–1450: Texts and Contexts, Essays and Studies*. Vol. 68. Woodbridge: Boydell and Brewer, 2015. 1–24.

De Bury, Richard. *The Philobiblon of Richard de Bury, Bishop of Durham, Treasurer and Chancellor of Edward III*, edited and translated by Ernest C. Thomas. New York: Lockwood and Coombes, 1889.

Derolez, Albert. *Codicologie des manuscrits en écriture humanistique sur parchemin*. Turnhout: Brepols, 1984.

DigiVatLib, https://digi.vatlib.it/mss/. Vatican Library.

Gillespie, Alexandra. "Books and Booklessness in Chaucer's England." In *The Oxford Handbook to Chaucer*, edited by Suzanne Conklin Akbari and James Simpson. Oxford: Oxford University Press, 2020. 81–97.

Grafton, Anthony. *Inky Fingers: The Making of Books in Early Modern Europe*. Cambridge: Harvard University Press, 2020.

Gumbert, J. P. "The sizes of manuscripts: some statistics and notes." In *Hellinga Festschrift, Feestbundel, mélanges: Forty-three studies in Bibliography presented to Prof. Dr. Wytze Hellinga on the Occasion of his retirement*, edited by A.R.A. Croiset van Uchelen. Amsterdam: Nico Israel, 1980. 277–88.

---. "Fifty Years of Codicology." *Archiv für Diplomatik* 50 (2004): 505–26.

Hanna, Ralph. "Miscellaneity and Vernacularity: Conditions of Literary Production in Late Medieval England." In *The Whole Book: Cultural Perspectives on the Medieval Miscellany*, edited by Stephen Nichols and Siegfried Wenzel. Ann Arbor: University of Michigan Press, 1996.

Jacob, Arun. "Punching Holes in the International Busa Machine Narrative." In *Alternative Historiographies of the Digital* Humanities, edited by Dorothy Kim and Adeline Koh, 121–44. New York: punctum books, 2021.

Johnston, Michael, and Michael Van Dussen, *The Medieval Manuscript Book*. Cambridge: Cambridge University Press, 2015.

Kestemont, Mike, et al. "Forgotten Books: The Application of Unseen Species Models to the Survival of Culture." *Science* 375.6582 (2022): 765–69.

Kim, Dorothy. "Introduction." In *Alternative Historiographies of the Digital Humanities*, edited by Dorothy Kim and Adeline Koh, 15–32. New York: punctum books, 2021.

---. "Embodying the Database: Race, Gender, and Social Justice." In *Alternative Historiographies of the Digital Humanities*, ed. Dorothy Kim and Adeline Koh, 145–202. New York: punctum books, 2021.

Kwakkel, Erik. "Biting, Kissing and the Treatment of Feet: The Transitional Script of the Long Twelfth Century." In *Turning Over a New Leaf: Change and Development in the Medieval Book*, edited by Erik Kwakkel, Rosamond KcKitterick, and Rodney Thomson. Leiden: Leiden University Press, 2012. 79–126.

Liu, Yin. "Ways of Reading, Models for Text, and the Usefulness of Dead People." *Scholarly and Research Communication* 5 (2015): 1–13.

Magnani, Roberta, and Diane Watt. "Towards a Queer Philology." *postmedieval: a journal of medieval cultural studies* 9 (2018), 252–68.

Nafde, Aditi, and Emma Gorst. "Introduction: The Trials of the Digital Medievalist." *Digital Philology* 4 (2015): 147–59.

Ornato, Ezio. *La face cachée du livre médiéval: l'histoire du livre vue par Ezio Ornato, ses amis et ses collègues*. Rome: Viella, 1997.

Ozment, Kate. "Rationale for Feminist Bibliography." *Textual Cultures* 13 (2020), 149–78.

Pietsch, Wolfgang. *Big Data*, Elements in the Philosophy of Science. Cambridge: Cambridge University Press, 2021.
Sargan, J. D. "What Could a Trans Book History Look Like? Towards Trans Codicology." *Criticism* 64 (2022): 571–86.
Sawyer, Daniel. *Reading English Verse in Manuscript, c.1350–c.1500.* Oxford: Oxford University Press, 2020.
Scott, Kathleen. *Dated and Datable English Manuscript Borders, c. 1395–1499*. London: Bibliographical Society, 2002.
Scott-Fleming, Sonia. *The Analysis of Pen Flourishing in Thirteenth-Century Manuscripts*. Leiden: Brill, 1989.
Spires, Derrick. "On Liberation Bibliography: The 2021 BSA Annual Meeting Keynote." *Papers of the Bibliographical Society of America* 116 (2022), 1–20.
Singh, Devani. *Chaucer's Early Modern Readers: Reception in Print and Manuscript*. Cambridge: Cambridge University Press, 2023.
Szirmai, J.A. *The Archaeology of Medieval Bookbinding*. Aldershot, UK: Ashgate, 1999.
Treharne, Elaine, *Perceptions of Medieval Manuscripts: The Phenomenal Book*. Oxford: Oxford University Press, 2022.
Underwood, Ted. *Distant Horizons: Digital Evidence and Literary Change*. Chicago: University of Chicago Press, 2019.
Warren, Michelle, and Neil Weijer, *Reimagining History: A Browsable Directory of the Middle English Prose Brut Chronicle*, Dartmouth College, https://www.library.dartmouth.edu/digital/digital-collections/brut-chronicle/handlist.
Weller, Susie, et al. *Big Qual: A Guide to Breadth-and-Depth Analysis*. Palgrave Macmillan, 2023.