

HISTORY *of the* BOOK

Afterword: Digitization

Digitization takes several forms and serves purposes of access, analysis, and original production of born-digital content as well as platforms and tools. In the humanities, and in information professions, such work has precedents in the visionary activity of early 20th century pioneer like Paul Otlet and Vannevar Bush.¹ But also, the science fiction visionaries H.G. Wells and George Orwell imagined a world in which print and manuscript materials would be remediated into electronic environments.² The role of the screen as a point of interface was crucial in all of their imaginings, and though it remains the main way of working with networked materials for now, it may not in the future as immersive and projective technologies gain ground. But if we take each of the three areas of access, analysis, and original production in turn, we can begin to see the ways that digitization extends literacy technologies—sometimes in imitation of older modes and sometimes innovatively. Sustainability and preservation are essential for working with digital assets and their management, but these topics that will not be addressed in this brief overview. Clearly preservation is as crucial to digital projects as it is to the care of works on vellum, clay, papyrus, or paper discussed in the previous chapters.

Access

The creation of digital facsimiles of analog material through photography or scanning is a common mode of migrating cultural legacy materials into digital format. The first generation of projects in which digitization took place saw the advantages of being able to bring together materials—manuscripts, rare books, unique editions or copies of published works, photographs and other unique items—that were geographically distributed. The William Blake and Rossetti Archives, two early projects, both aimed at comprehensive representation and study of the works of these individual artists.³ In the first instance, the research questions focused on technique, variation, and the effect of individual illumination on meaning production in Blake's artistic books. The Rossetti project sought to extend the traditions of critical editing and bibliographical scholarship by being able to collect, display,

and annotate the variorum editions of all of Rossetti's works. Such projects, not realizable in an actual space—the institutional repositories would never have released the materials for such a purpose—were demonstrations of the potential that digital environments had for supporting study of books and their history.⁴

Access drove (and continues to drive) many large-scale projects. The migration of major treasures into online format has included such works as the Book of Kells, the Codex Sinaiticus, the Homeric textual witnesses (earliest manuscripts), Shakespeare first folios, manuscripts of Beowulf, and so on. In addition, collections of primary materials useful for the study of history, culture, politics, philosophy, music, art and other disciplines have been made available for research, classroom use, and general audiences. This has made these materials usable wherever internet access reaches. While this is not a fully global range, it brings the materials to the users, helping to make access more equitable than it would be if the only way to see these objects were by travelling to where they are stored. Wear and tear is greatly reduced by having digital facsimiles available for study. In addition, multiple users can access the same document at the same time—and compare it with other materials that are stored somewhere far away. Scanning resolution is sufficiently high that more information is sometimes available on screen than to the naked eye, and in addition, commentary and authoritative annotation are also often available. While digitization is not a panacea for inequitable distribution of resources, it has the benefit of at least making primary resources available more widely than at any prior moment in history.

Access and use are not the same, of course, and even if materials exist in digital form, the skills for engagement and the degree of access vary considerably. The digitization of rare materials increases the need for interpretation and frameworks within which these objects can be received. The issues that arise around indigenous materials, sensitive documents, and other items to which access might need to be selectively managed, are also ones that touch on ethical and political matters. The professional communities of archivists, librarians, scholars, and curators take such factors into account in thinking about their role in managing cultural artifacts.

Analysis

Once textual, visual, audio, or spatial material is made computationally tractable, it is available for manipulation of various kinds. The use of analytic tools for data mining and textual analysis, among other tasks, has produced new kinds of artifacts for research in the history of books and written remains. The use of various imaging technologies, such as MRI and infrared, though familiar from their use in revealing features of analogue materials not available to the naked eye, is amplified when the distinctive elements each imaging can discern are computationally aggregated. A collection of ostraca, found in a fort

in Israel, were subjected to scanning and sampling of their materials and analysis of the handwriting on the small shards in a project whose findings were published in 2016.⁵ The texts were mainly dealing with mundane matters such as ordering supplies and managing inventory, but by assessing the number of hands involved, scholars were able to extrapolate an estimate of literacy rates in a population that could have been contemporary with early Biblical composition. These discoveries allow fragmentary historical evidence to be analyzed with far greater detail and sophistication than through traditional paleography and archaeology. The complement of traditional and digital methods comes to the fore, and though analytic studies of large textual and visual corpora show patterns in the material structure of these objects, the question of how to integrate such literal analysis into interpretative techniques around substantive questions of meaning, value, cultural identity and humanistic value has not been fully addressed.

For historical investigation, textual analysis and data mining are an aid to search and query, to the discovery of patterns within bodies of materials that are too large for a single reader to analyze.⁶ Because analysis can be performed on structured data (taxonomies, classification, metadata, texts marked in XML or other interpretative schemes) as well as unstructured data (raw texts, images, sound files), it can point to places from which scholarship might begin. The notion of literacy is affected by these techniques, which perform their analytics within a set of algorithmically driven procedures that search, count, and match character strings. These do not allow for the sustained ambiguity and contradiction that are central to certain traditions of scholarship, particularly in the humanities. Literacy extends to the requirement that a scholar, student, or even general reader be able to have a sense of how a digital analysis has been produced—what were the sources, how big is the sample, how are statistical errors and averages taken into account—if these artifacts are to be read responsibly. As these analytic tools increase in use, their black-box operations might need to be made transparent as part of the technology of digital literacy.

Original production

In addition to the burst of creative work done on electronic platforms at the end of the 20th century, the works of hyper-fiction and algorithmic, procedural, poetics, on which a few remarks were made in the last chapter, a whole host of innovations in publishing have been implemented that push the definition of a “book” and artifact into blurry domains. If a book was originally defined in the codex form, with its apparent boundedness and fix sequence of signatures, and distinguished from a scroll by virtue of the random access to its pages (not the linear access to which scroll forms are limited), then can the e-book, which scrolls in a relentlessly linear manner on the screen, be considered a cousin to the codex or the progeny of a scroll?⁷ And as electronic documents, with their contingent dependence on

conditions of production (clock speeds, processing speeds, software and hardware relations, bandwidth, networking protocols, and algorithms for display), continue to expand in use, how are they to be understood? Electronic documents lack the stability of analog ones, not because they are, as is often mistakenly stated, *immaterial*, but on the contrary, because they are so complexly *material*, as is evident by the description above.⁸ The fluidity and vulnerability of electronic texts is palpable, as the cycles of obsolescence make dramatically clear. Works of electronic text made twenty years ago may not be able to be viewed or engaged except in a simulator that reproduces the original conditions. Additional questions of reliability, authenticity, stability, and legitimacy are raised by such fluid and dynamic works—which can masquerade in multiple ways.

Publication models for electronic work continue to evolve, and the specific capacities of networked environments have yet to fully assume the conventions they will have ahead. These capacities include the ability to aggregate dispersed materials, produced interpretative frameworks for these aggregate projects, create arrays of search/retrieval and links for related works, corpora, and documents. A line of inquiry into an assembled corpus might follow its historical conditions, its political philosophy, its bibliographical history, its place within the life and work of an author, its impact and reception—all simultaneously. To find our way through these forms of argument, familiar conventions will have to emerge for the ways we track, read, assess, and preserve relations among documents, their analysis, their provenance, and interpretation. All of this will unfold ahead, as tools for digital scholarship and expertise in their design and use continue to emerge. Literacy will be, as it has been with the technologies of the past, a matter of combining familiarity with inquiry, trust with skepticism, and acts of reading with those of interpretation within cultural, historical, and theoretical frames.

* * *

Notes

¹ For discussion of Paul Otlet, see Ronald E. Day, *The Modern Invention of Information* (Carbondale, IL: Southern Illinois University Press, 2001) and for a general history of digital media, Martin Campbell-Kelly and William Aspray, *The Computer: History of the Information Machine* (NY: Harper Basic Books, 1996).

² H.G. Wells, *World Brain* (London: Methuen and Co., 1938) and George Orwell, *1984* (London: Harvill Secker, 1949).

³ The Rossetti Archive, <http://www.rossettiarchive.org/> (accessed 6/12/2017) and The William Blake Archive, <http://www.blakearchive.org/> (accessed 6/12/2017).

⁴ Ray Siemens and Susan Schreibman, *A Companion to Digital Literary Studies* (Oxford: Blackwell, 2008) and Jerome McGann *Radiant Textuality: Literary Studies After the World Wide Web* (London and NY: Palgrave, 2001).

⁵ Ostraca in Israel read using new technology; Amanda Borschel-Dan, “Revolutionary Technology Reveals Dazzling Hidden Text on Biblical Era

Shard,” <http://www.timesofisrael.com/revolutionary-technology-reveals-dazzling-hidden-text-on-biblical-era-shard/> (6/17/2017) and earlier, by Sharona Schwartz, “Ancient Pottery Shards Analyzed,” <http://www.theblaze.com/news/2015/04/22/ancient-pottery-shards-analyzed-by-israeli-scientists-seem-to-support-biblical-narrative/>

⁶ Stéfan Sinclair and Geoffrey Rockwell, *Hermeneutica* (Cambridge, MA: MIT Press, 2016).

⁷ Johanna Drucker, “From page space to e-space,” *Philobiblon*. April 25, 2003; <http://www.philobiblon.com/drucker/>

⁸ Jean-François Blanchette, “A Material History of Bits,” *JAIST: Journal of the Association for Information Science and Technology*, April 20, 2011; <http://onlinelibrary.wiley.com/doi/10.1002/asi.21542/abstract> (6/17/2017).