Visualization and Digital Collections

Since the 1990s, cultural heritage institutions have been investing in digital technologies to address a growing public demand for open and permanent access to information resources. Accordingly, galleries, libraries, archives, and museums worldwide have strategically focused on the digitization of their holdings. The next step involved the development of digital collections and services in support of online research and learning.[1]

While enabling direct access to cultural and scientific heritage, digitization of archival materials has also fostered their preservation, virtual collocation on web portals, and the creation of an integrated learning environment. As a result, libraries have seen a substantial increase in the use of digitized materials that have attracted diverse users.[2] Researchers, scholars, educators, entrepreneurs, and web surfers have engaged with content on an unprecedented scale because it is available in digital format.[3]

This digital shift in the library world continues to accelerate. Due to the pandemic crisis in 2020, print collections have rapidly become unavailable. Research and learning have moved to a virtual environment for the immediate future, and perhaps for good. Digital content has suddenly transitioned from being a preview of the physical collection to its primary access point.[4]

Digital collections, however, are not simply representations of physical collections. Rather, they are resources in their own right.[5] Unlike physical collections, digital ones have detailed metadata and, often, full text available due to OCR conversion of text images into machine-encoded data. Both metadata and data can be mined, analyzed, and visualized. *Text mining* refers to discovering and extracting meaningful patterns from large numbers of text documents. Practitioners of computational linguistics and digital humanities call these agglomerations of texts *corpora* or *capta*.[6] Text analysis of *corpora* or *capta* at its very basic level involves looking at words' frequencies, contexts, lexical preferences, and relations among textual elements. *Visualization* refers to the design of the graphic representations of data objects and their relationships.[7] It is commonly viewed as an essential part of textual data analysis that helps to make quantitative information legible and easy to comprehend.[8]

Digital collections' data are open for exploration and analysis just as any other data are. Similarly to humanists, who now examine how technology is changing our understanding of the liberal arts, we can employ digital tools to see how they may shape our understanding of archival curation and librarianship.[9] We begin with examining some pragmatic reasons for visualization of digital collections. Then our focus will move to the particular application of visualization tools, specifically the R programming language, including ggpolt2 and RShiny. R has been designed to facilitate data analysis, statistics, statistical programming, and graphics.[10]

Why Visualize a Digital Collection's Data

There are multiple reasons to visualize a digital collection's data, but one of the most important is to appeal to our natural visual abilities. Humans are very good at seeing patterns. The biggest part of our brain cortex is involved in visual perception.[11] According to the Light Switch Theory by Andrew Parker, the eyes and vision have been the primal driving force of biological evolution since the Cambrian Period, approximately 543 million years ago.[12] The rapid development of visual perception has thus been considered a fundamental survival strategy for animal species, including humans. Tuned by evolution, human vision has always been our most critical sense for detecting and extracting information from our physical surroundings.[13]

Naturally, various forms of visualization, as means of communication utilized by people, have been

present worldwide throughout human history. Cave paintings and rock art, maps, calendars, genealogical trees, chronological tables, and time lines, as well as modern-day charts and graphs, have all been used to convey essential information so that it can be grasped quickly.[14] This high efficiency of an image that "is worth a thousand words" has recently played a leading role in the development of present-day screen culture. In his book on this global phenomenon, Richard Butsch formally defines *screen culture* and points to its continuing relevance:

> Screens are about images more than language, a modern form of visual culture. . . . It is the *lived* culture that arises when people interact with and through screen media. As everyday life fills with screen activity—averaging nine hours per day for American adults in 2017—screen becomes an increasingly important aspect of the broader culture infiltrating and influencing all other elements.[15]

Indeed, visualization presently seems to dominate all communications.[16] Films, television, video games, video websites, mobile applications, and social media circulate images on a massive scale. We interact with dashboards, maps, newspaper charts, app graphs, and text messages enhanced by emojis or animated GIFs on an hourly basis. Visual representations have become an integral part of our social and professional communication practices. Scott Berinato refers to visualization as a lingua franca used in knowledge economies of the twenty-first century.[17] He argues that the attractiveness of visualization stems from its ability to summarize and simplify complex data and thus help to make sense of them. In addition, the practice of data visualization has now been demystified and democratized.[18] It is no longer in the purview of only software experts and professional coders. Due to technological advances and growing demand, tools for creating graphs and charts have become open and available to everyone who is willing to learn a new language.[19] If libraries are to participate and contribute to screen culture, then perhaps it is time to learn its language and go visual.

There is a growing literature on the relevance of graphics for digital libraries. Visualization is often used to support exploration and discovery, content analysis, and communication about the collection. For example, graphical representations of digital collections are considered to be a great alternative to text-based interfaces and search boxes, especially for non-experts and casual users.[20] Unlike empty search fields that rely on the users' input and background knowledge, graphs and diagrams provide a comprehensive overview of the collection easily understandable by all users. Along the same lines, "generous interfaces" are designed to show graphs of digital collections up front on web portals in order to both spark users' interest and inspire further exploration of digitized material.[21] In addition to providing an overview of a collection's scope and content, generous interfaces include the contexts for the collection, a display of the relationships among collection items, and a quick closer look at selected images. These graphic overviews are natural starting points for browsing large sets of digital items, identifying relevant topics and patterns, selecting pertinent documents and images, and finally focusing on their details.[22] Recent implementations of visual search and discovery involve interactive interfaces where users navigate digital collections as virtual galleries.[23] According to Eric Phettelace, interactivity adds an additional discovery layer that enables users to become active agents in finding new patterns in data and putting new interpretations on them.[24]

Indeed, graphics foster not only intended searches for information, but also serendipitous findings. According to Windhager and colleagues, following graphics and diagrams often leads users to discoveries of diverse perspectives about a collection's data.[25] These unexpected findings, in turn, may inspire new approaches to examination of historical evidence and new ways of thinking about the nature of primary sources and information at large. Since charts and diagrams may easily reflect a diversity of views and the complexity of information, they seem to be very well suited for multithreaded investigation.[26]

Similarly, archivists and curators find the application of graphics extremely useful for the analysis of large digital collections.[27] Visualization allows curators to examine the structure and organization of a collection, its content and provenance, relationships among the collection's items, the scope and size of the collection, and the number of files and their formats, as well as text patterns in documents and visual patterns among images. In addition, graphs may reveal distributions of various documents and images over time that provide remarkable insight into the process of collection development. Monitoring progress of a collection also involves the assessment of its metadata in terms of their completeness and quality. Computing applications used for visualization fully expose all inconsistencies and missing values across metadata fields. Visualization then may also be used as an effective tool for metadata quality control. All these observations about a collection's structure and a description of its content come with a growing understanding of data enabled by visualization. Richard Hamming once pointed out that the purpose of computing is insight—not numbers.[28] By the same token, the purpose of visualization is grasp rather than graphics.

To enable readers to get a good grasp on data, graphics need to communicate information clearly. Graphs and diagrams are to explain data, not to

obscure them.[29] According to Richie Cotton, the effectiveness of a plot can be measured by two criteria: how many insights readers can get from the plot and how fast they can do so.[30] Some plots are specifically designed to tell stories. There seems to be a strong relationship between data, visualization, and narrative, especially when it comes to graphic representation of time and chronology.[31] The concept of mapping duration to space, and particularly to the length of a line, was first put into practice by Joseph Priestly in 1764 to compare the life spans of 2,000 famous people.[32] The idea of time represented by a measurable line has evolved ever since, reflected in drawn streams, chains, trees, and now time lines. Time has become an organizing principle for diagrams. But it is also the organizing principle for narratives. Like graphs, narratives represent events in time.[33] Like graphs, stories help people see.

Digital data have their own stories to tell.[34] They also have their own ways of doing so. Traditionally, graphs and diagrams have been used to support narratives with envisioned concrete evidence embedded in a text sequence. This sequence usually starts with an overview of data in order to set the scene and context for their interpretation. After setting the opening scene, the creator of a graph guides the readers through its most prominent features and smoothly directs users' attention from one point of interest to the next one. Less prominent data features, which do not add to the main story, are typically left out, as the author tries to present the data in the most convincing way to advance the line of a given argument. However, the full complexity of digital data may not be adequately represented by one leading story line.[35] Rather, it calls for multiple interpretations depending on readers' interests. Once again interactivity of graphs seems to invite and enable users to create their own story lines and paths of discovery. The plots developed by users may diverge considerably from the order suggested by authors and follow various unexpected directions. Users may remix data and reinvent the entire interpretation of a collection. A combination of a traditional narrative approach and interactive elements that foster user-driven exploration is becoming a standard in designing visual representations.

Indeed, visual communication tends to work best when the audience gets engaged in the communication process.[36] Clearly, active learning radically improves comprehension of data. It also awakens users' curiosity. The key objective for developing visualizations of digital collections should then be to inspire and actively engage users with digital content. But the users need to be familiar with visual interfaces and feel comfortable interacting with them in the first place.[37]

The GLAM Labs community provides support for users seeking to feel more comfortable interacting with visualizations.[38] Galleries, libraries, archives, and museums (GLAM) have promoted digital content for reuse and experimentation in educational, commercial, and artistic projects. For the GLAM Labs community, digital cultural heritage is not just to contemplate, but also to fully engage with in creative ways. Accordingly, Europeana Pro, the British Library Digital Scholarship Department, the Library of Congress's LC Labs, and the Digital Public Library of America—DPLA Pro all now offer comprehensive guidelines for how to work and innovate with digital collections.[39]

The development of digital cultural heritage agglomerations, like the Europeana Project that started in 2007 and the DPLA established in 2010, have paralleled the rise of the digital humanities (DH) as a new research field with its own questions and methodology. Digital collections are primary research sources for DH scholars, whereas data visualization is one of their essential methods for text and data analysis.[40] Anne Burdick and her associates argue that visualization in fact provides "graphical legibility to analytical results."[41] In their view, geo-temporal visualizations and mapping allow scholars to examine complex interrelations among cultural, social, and historical phenomena.

In contrast with a traditional humanities research approach that emphasizes individual authorship, a digital approach fosters cooperation and partnership. Based on a survey of five hundred scholars, librarians, and archivists, Jessica Wagner Webster shows that there are multiple opportunities for successful collaboration among these stakeholders in regard to digital projects.[42] Interestingly, the roles that these stakeholders play do not always align with expected tasks, in which DH scholars come up with research questions and interpret the results, while librarians and archivists are in charge of digital tools and their implementation. One of the benefits of interdisciplinary projects is a variety of views brought together. This is exactly where innovation begins.

A Process for Getting to Know Data

As discussed earlier, graphic representation is among our most important tools for organizing data and sharing information. The process of creating effective visualizations of given data involves some important preliminary steps, including selecting visualization tools and preprocessing data compiled in a table. The first step of this preprocessing is examining the data table and getting familiar with collection data and metadata. Gaining detailed knowledge about data allows consideration of which aspects of information contained in the data might be of interest to collection users or curators.

One of the tools that may help a user to get a clear insight about collection data is OpenRefine.[43] It is a free, open source software application for working with raw and messy data. OpenRefine allows for importing data in various formats, exploring large data sets, cleaning and transforming data, and also linking data sets with web services: for instance, getting geographic coordinates for addresses. OpenRefine runs on all major operating systems, including Windows, macOs, and Linux.

An OpenRefine project operates similarly to a spreadsheet or a table consisting of columns with metadata elements and rows of data. The rows can be filtered by various criteria and can also be edited. OpenRefine allows for detailed examination of the collection content and its description.

After examining data, the next step involves developing specific questions about a digital collection that compiled data may potentially address. In fact, these questions inform the initial mining of raw data and metadata. For this reason, this is the stage where interdisciplinary collaboration is particularly relevant because it brings diverse interests and questions together. Depending on the questions asked, the relevant pieces of information are extracted from the data table. These selected pieces are then closely examined for completeness and consistency. By nature all data, including metadata, tend to be messy. Therefore cleaning or tidying data is an essential prerequisite for their effective visual display.

Following data cleanup, the next step, often necessary, is transforming data. Data transformation allows for obtaining defined values necessary for plotting and graphic display. For example, extracting time measures in specific units from string date representations, aggregating data points according to different categories, and applying mathematical formulas to column values may be needed to obtain well-defined sets of specific numbers. The graphs are created by subjecting clean and transformed data to plotting functions that translate numbers into their graphic representations. The final step in the visualization process is tuning graphs for clarity.

As mentioned earlier, the number of tools for visualization is continually increasing. Many of them, including IBM Many Eyes, Library of Congress Viewshare, Microsoft Excel, Tableau, D3.js, FusionCharts, Google Charts, Dygraphs, Infogram, Plotly, IBM Watson Analytics, Tableau Public, TimelineJS, StorymapJS, Google Maps, and Historypin are discussed elsewhere.[44]

This report focuses on the application of the R programming language and its specific packages—ggplot2 for visualization and RShiny for interactive visualization. The next chapter addresses the methodology of learning the R programming language, the general workflow for basic visualization, and an introduction to graphic representations of data that serve specific analytical tasks.

Acknowledgments

Notes

1. Daniel Greenstein, "Digital Libraries and Their Challenges," *Library Trends* 49, no. 4 (2000): 290–303; Laura Deal, "Visualizing Digital Collections," *Technical Services Quarterly* 32 (2015):14–34; "IFLA/UNESCO Manifesto for Digital Libraries," International Federation of Library Associations and Institutions, accessed May 20, 2020, https://www.ifla.org/publications/iflaunesco-manifesto-for-digital-libraries; Florian Windhager, Paolo Federico, Günther Schreder, Katrin Glinka, Marian Dörk, Silvia Miksch, and Eva Mayr, "Visualization of Cultural Heritage Collection Data: State of the Art and Future Challenges," *IEEE Transactions on Visualization and Computer Graphics* 25, no. 6 (2019): 2311–2330, https://doi.org/10.1109/TVCG .2018.2830759.

2. Peter B. Hirtle, "The Impact of Digitization on Special Collections in Libraries," *Libraries & Culture: A Journal of Library History* 37, no. 1 (2002): 42–52, https://hdl.handle.net/1813/14206.

3. M. Mahey, A. Al-Abdulla, S. Ames, P. Bray, G. Candela, S. Chambers, C. Derven, et al., *Open a GLAM Lab* (Doha, Qatar: Qatar University Press, 2019), https://glamlabs.pubpub.org/.

4. Christopher Cox, "Changed, Changed Utterly," *Inside Higher Ed,* #Views, #Opinion, June 5, 2020, https://www.insidehighered.com/views/2020/06/05 /academic-libraries-will-change-significant-ways -result-pandemic-opinion.

5. Florian Kräutli, "Visualising Cultural Data: Exploring Digital Collections through Timeline Visualisations" (doctoral thesis, Royal College of Art, London, 2016), http://www.kraeutli.com/index.php/2016/04/15 /visualising-cultural-data/.

6. David Crystal, *The Cambridge Encyclopedia of Language*, 3rd ed. (New York: Cambridge University Press: 2010), 434–435;Johanna Drucker, "Digital Humanities: Approaches to Graphical Display," *DHQ: Digital Humanities Quarterly* 5, no. 11 (2011), http://www .digitalhumanities.org/dhq/vol/5/1/000091/000 091.html; Anne Burdick, Johanna Drucker, Peter Lunenfeld, Todd Presner, and Jeffrey Schnapp, *Digital Humanities* (Cambridge, MA: MIT Press, 2012).

7. Katy Börner, Chaomei Chen, and Kevin W. Boyack, "Visualizing Knowledge Domains," *Annual Review of Information Science and Technology* 37, no. 1 (2003):

179–255, https://doi.org/10.1002/aris.1440370106.

8. Johanna Drucker, "5B. Data Mining and Text Analysis," in Johanna Drucker with David Kim, Iman Salehian, and Anthony Bushong, *Introduction to Digital Humanities: Concepts, Methods, and Tutorials for Students and Instructors*, course book for DH 101: Intro to Digital Humanities (Los Angeles, CA: UCLA Center for Digital Humanities, 2014), 43–45, http://dh101.humanities.ucla.edu/?page_id=48; Scott Berinato, *Good Charts: The HRB Guide to Making Smarter, More Persuasive Data Visualizations*, (Boston: Harvard Business Review Press, 2016).

9. Patricia Cohen, "Humanities 2.0: Digital Keys for Unlocking the Humanities' Riches," *New York Times*, November 16, 2010, https://www.nytimes.com/2010/11/17/arts/17digital.html?auth=link-dismiss-google1tap.

10. Hadley Wickham and Garrett Grolemund, *R for Data Science: Import, Tidy, Transform, Visualize and Model Data* (Sebastopol, CA: O'Reilly Media, 2017); J. D. Long and Paul Teetor, *R Cookbook: Proven Recipes for Data Analysis, Statistics and Graphics* (Sebastopol, CA: O'Reilly Media, Inc., 2019).

11. Margaret S. Livingstone, *Vision and Art: The Biology of Seeing*, rev. and exp. ed. (New York: Abrams, 2014).

12. Andrew Parker, *In the Blink of an Eye: How Vision Sparked the Big Bang of Evolution* (Cambridge, MA: Perseus, 2003).

13. Livingstone, *Vision and Art*; William S. Cleveland and Robert McGill, "Graphical Perception: Theory Experimentation and Application to the Development of Graphical Methods," *Journal of the American Statistical Association* 79, no. 387 (1984): 531–54; William S. Cleveland and Robert McGill, "Graphical Perception and Graphical Methods for Analyzing Scientific Data," *Science* 229 (August 30, 1985): 828–33.

14. Michael Friendly, "Milestone in the History of Data Visualization: A Case Study," in *Classification: The Ubiquitous Challenge: Proceedings of the 28th Annual Conference of the Gesellschaft für Klassifikation e.V., University of Dortmund, March 9–11, 2004*, ed. Claus Weihs and Wolfgang Gaul, 34–52 (Heidelberg, Germany: Springer-Verlag, 2005); Daniel Rosenberg and Anthony Grafton, *Cartographies of Time: A History of the Timeline* (New York: Princeton Architectural Press, 2010); Stephen Boyd Davis, Olivia Vane, and Florian Kräutli, "Using Data Visualisation to Tell Stories about Collections" (paper, Electronic Visualisation and the Arts [EVA] conference, London, UK, July 12–14, 2016), https://doi.org/10.14236/ewic/EVA2016.44; Berinato, *Good Charts*; Hsuanwei Michelle Chen, "Information Visualization," *Library Technology Reports* 53, no. 3 (April 2017).

15. Richard Butsch, *Screen Culture: A Global History* (Cambridge, UK Polity Press, 2019): 2–3.

16. Berinato, *Good Charts*.

17. Berinato, *Good Charts*.

18. Eric Phettelace, "Effectively Visualizing Library Data," *Reference and User Services Quarterly* 52, no. 2 (2012): 93–97.

19. Jannette L. Finch and Angela R. Flenner, "Using Data Visualization to Examine an Academic Library Collection," *College and Research Libraries* 77, no. 6 (2016): 765, https://doi.org/10.5860/crl.77.6.765; Chen, "Information Visualization;" Kayla Harris and Andrew Harris, "Data Visualization Tools for Archives and Special Collections," *MAC Newsletter*, Midwest Archives Conference, 2019: 26–29, https://ecommons.udayton.edu/imri_faculty_publications/39.

20. Panayiotis Zaphiris, Kulvinder Gill, Terry H. Y Ma, Stephanie Wilson, and Helen Petrie, "Exploring the Use of Information Visualization for Digital Libraries," *New Review of Information Networking* 10, no. 1 (2004): 51–69, https://doi.org/10.1080/13614570420000304136; Ali Shiri, "Metadata-Enhanced Visual Interfaces to Digital Libraries," *Journal of Information Science* 34, no. 6 (2008): 763–75, https://doi.org/10.1177/0165551507087711; Mitchell Whitelaw, "Towards Generous Interfaces for Archival Collections," (paper presented at the International Council on Archives Congress, Brisbane, Australia, August 20–24, 2012), http://mtchl.net/assets/Whitelaw_ICA_GenerousInterfaces.pdf; Deal, "Visualizing Digital Collections"; Florian Windhager, Paolo Federico, Eva Mayr, Günther Schreder and Michael Smuc, "A Review of Information Visualization Approaches and Interfaces to Digital Cultural Heritage Collections," in *Proceedings of the 9th Forum Media Technology 2016 and 2nd All Around Audio Symposium (FMT 2016), St. Pölten, Austria, November 23–24, 2016*, ed. Wolfgang Aigner, Grischa Schmiedl, Kerstin Blumenstein, Matthias Zeppelzauer, and Michael Iber, *CEUR Workshop Proceedings*, http://ceur-ws.org/Vol-1734/fmt-proceedings-2016-paper9.pdf; Windhager et al., "Visualization of Cultural Heritage Collection Data."

21. Whitelaw, "Towards Generous Interfaces."

22. Zaphiris et al., "Exploring the Use of Information"; Mark Hall and Paul Clough, "Exploring Large Digital Library Collections Using a Map-Based Visualisation," in *Research and Advanced Technology for Digital Libraries: International Conference on Theory and Practice of Digital Libraries, TPDL 2013, Valetta, Malta, September 22–26, 2013, Proceedings,* ed. Trond Aalberg, Christos Papatheodorou, Milena Dobreva, Giannis Tsakonas, and Charles J. Farrugia, 216–27, *Lecture Notes in Computer Science* 8092, (Berlin, Heidelberg: Springer, 2013); Deal, "Visualizing Digital Collections"; I. di Lenardo, B. Seguin, and F. Kaplan, "Visual Patterns Discovery in Large Databases of Paintings." Digital Humanities 2016: Conference Abstracts (Kraków: Jagiellonian University and Pedagogical University, 2016): 169–72, https://dh2016.adho.org/abstracts/348.

23. Jeffrey P. Emanuel, Christopher M. Morse, and Luke Hollis, "The New Interactive: Reimagining Visual Collections as Immersive Environments," *VRA Bulletin* 43, no. 2, (2016), article 2, https://online.vraweb.org/index.php/vrab/article/view/52/57.

24. Phettelace, "Effectively Visualizing Library Data."

25. Windhager et al., "Visualization of Cultural Heritage Collection Data."

26. Shiri, "Metadata-Enhanced Visual Interfaces"; Kräutli, "Visualising Cultural Data"; Anne Bahde, "Conceptual Data Visualization in Archival Finding Aids: Preliminary User Responses," *portal: Libraries and the Academy* 17, no. 3 (2017): 485–506; A. Miller, "Data Visualization as Participatory Research: A Model for Digital Collections to Inspire User-Driven Research,"

Journal of Web Librarianship 13, no. 2 (2019): 127–77, https://doi.org/10.1080/19322909.2019.1586617.

27. Shiri, "Metadata-Enhanced Visual Interfaces"; Weijia Xu, Maria Esteva, Suyog Dutt Jain, and Varun Jain, "Analysis of Large Digital Collections with Interactive Visualization," in *2011 IEEE Conference on Visual Analytics Science and Technology (VAST)* (New York: IEEE, 2011), 241–50, https://doi.org/10.1109/VAST.2011.6102462; Kräutli, "Visualising Cultural Data."

28. Richard W. Hamming, *Numerical Methods for Scientists and Engineers* (New York: McGraw-Hill, 1962), vii, 276, 395.

29. Darrell Huff, *How to Lie with Statistics* (New York: W. W. Norton, 1954); Ray Lyons, "Beauty Is as Beauty Does," *Lib(rary) Performance* (blog), October 28, 2011, https://libperform.wordpress.com/2011/10/28/beauty-is-as-beauty-does/.

30. Richie Cotton, "Data Visualization for Everyone: An Introduction to Data Visualization with No Coding Involved," course on DataCamp platform, accessed June 26, 2020, https://learn.datacamp.com/courses/data-visualization-for-everyone (requires sign in).

31. Davis, Vane, and Kräutli, "Using Data Visualisation"; Kräutli, "Visualising Cultural Data."

32. Kräutli, "Visualising Cultural Data."

33. H. Porter Abbott, *The Cambridge Introduction to Narrative*, 2nd ed. (Cambridge: Cambridge University Press, 2008).

34. Edward Segel and Jeffrey Heer, "Narrative Visualization: Telling Stories with Data," *IEEE Transactions on Visualization and Computer Graphics* 16, no. 6 (2010), http://vis.stanford.edu/papers/narrative.

35. Wita Wojtkowski and W. Gregory Wojtkowski, "Storytelling: Its Role in Information Visualization" (paper, European Systems Science Congress, Crete, Greece, October 16–19, 2002), http://citeseerx.ist.psu.edu/viewdoc/citations;jsessionid=84B03925EA0F41B00DA79F9EE03F5932?doi=10.1.1.99.4771.

36. Yea-Seul Kim, Katharina Reinecke, and Jessica Hullman, "Explaining the Gap: Visualizing One's Predictions Improves Recall and Comprehension of Data," in *CHI '17: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (New York: Association for Computing Machinery, 2017), 1375–86, https://doi.org/10.1145/3025453.3025592.

37. Deal, "Visualizing Digital Collections"; Windhager et al., "A Review of Information Visualization Approaches"; Whitelaw, "Towards Generous Interfaces."

38. "Introducing GLAM Labs," in M. Mahey, A. Al-Abdulla, S. Ames, P. Bray, G. Candela, S. Chambers, C. Derven, et al., *Open a GLAM Lab* (Doha, Qatar: Qatar University Press, 2019), 33–45, https://glamlabs.pubpub.org/pub/kly6t0jq.

39. Europeana Pro home page, accessed May 12, 2020, https://pro.europeana.eu/; "Digital Scholarship," British Library, accessed June 2, 2020, https://www.bl.uk/subjects/digital-scholarship; "LC Labs," Library of Congress, accessed June 2, 2020, https://labs.loc.gov/; "DPLA Pro," Digital Public Library of America, accessed May 29, 2020, https://pro.dp.la/.

40. Burdick et al., *Digital Humanities*; Drucker, "5b. Data Mining and Text Analysis."

41. Burdick et al., *Digital Humanities*, 18.

42. Jessica Wagner Webster, "Digital Collaborations: A Survey Analysis of Digital Humanities Partnerships between Librarians and Other Academics," *DHQ: Digital Humanities Quarterly* 13, no. 4 (2019), http://www.digitalhumanities.org/dhq/vol/13/4/000441/000441.html.

43. OpenRefine home page, accessed May 24, 2020, https://openrefine.org/.

44. Phettelace, "Effectively Visualizing Library Data"; Deal, "Visualizing Digital Collections"; Finch and Flenner, "Using Data Visualization"; Chen, "Information Visualization"; Harris and Harris, "Data Visualization Tools."