Beyond Access: Critical Catalog Constructions

Molly Hardy
mhardy@mwa.org
American Antiquarian Society, United States of America

Dawn Childress
dchildress@library.ucla.edu
UC Los Angeles, United States of America

Paige Morgan
paige.c.morgan@gmail.com
University of Miami, United States of America

This panel explores the use of digital, rare book catalogs as platforms for collaboration and as sources of data to uncover patterns of book production and to offer new insights into the sociology of texts. Extractions from and additions to bibliographic data extend the catalog beyond its original use as a point of access and discovery. The use of catalogs as sources of bibliographic big data as well as the use of platforms that enable bibliographic record annotation allow us to reimagine the catalog's trajectories of creation and utility. Considering the genealogy of the catalog, the panel will examine how research practices have both informed and been informed by catalogs since their inception and how the work of bibliography is reimagined in the catalog. Focusing on how limitations and aporia in the catalog can lead to critical making in the digital age, the panel will show how digital uses of the catalog currently enable mindful interrogation of catalog data and catalog making as well as consider possibilities for expanded use of rare book catalog data in the future. This consideration of the rare book catalog as a digital humanities project invites reassessment of legacy information architecture as well as the many hands that built the bibliographic structures on which so much of the work of the digital humanities rests.

In "Towards Speculative Catalogs," Dawn Childress will open with a discussion of the transformative promise of the digital as we reconstruct catalogs in new forms and formats. To provide context for critical catalog constructions more broadly, Childress will highlight how bibliographies and catalogs have served as source material for research beyond that of points of access in both pre- and post-DH contexts, as well as consider how digital humanities use cases might differ from more analog approaches, whether qualitative or quantitative. In addressing these questions, Childress will explore the promise of applying current and emerging tools and standards (such as linked data, IIIF, PCDM, etc.) to the practice of interrogating bibliographic and catalog data. Childress will suggest how we might leverage these systems to record and analyze lacunae, erasures, and bias in capturing the bibliographic record and, drawing on Bethany Nowviskie's notion of speculative collections, how these systems might support active reframing and interrogation by users.

In "'The Technology of Shared Cataloging': A Retrospective," Molly O'Hagan Hardy will build on Childress's remarks through a close look at the creation and re-creation of two rare book union catalogs: the English Short Title Catalog (ESTC) and the North American Imprints Program (NAIP). In 1981, in a Bibliographic Society of America Symposium from which the title of Hardy's paper takes its name, William Todd wrote, "Perhaps we do not yet fully appreciate the situation, now rapidly materializing, whereby computers converse with each other in any mode, while the rest of us, mere mortals, stand mute before them." Remarks like this, which abound in the excitement and trepidation expressed during the emergence of these rare book union catalogs, echo a similar exuberance and hesitancy around the transformation from MARC to linked data models. Examining what "machine readable" meant then and means now, Hardy will draw parallels between the current conversation around BIBFRAME and other such initiatives and those early efforts led by Robin Alston and Marcus McCorison to amass large amounts of special collections' catalog data. She will then examine what it was these catalogs set out to capture and in what ways this work is being reimagined in the linked open data environment. She will do this through a close look at the American Antiquarian Society's Printers' File linked data project and its reliance on LCNAF and VIAF to merge MARC data with BIBFRAME. Ultimately, she will consider how such initiatives necessitate the reimagining of library and scholarly work, so those working on both sides of the reference desk are not left to "stand mute" before their creations. Hardy will point to examples of innovative uses of rare book catalog data for digital humanities projects. Such uses, Hardy will show, not

only prove efficient and effective means of generating data, but they also productively unearth biases inherent in any information system.

Paige Morgan will conclude our presentations with "Searching For Common Ground: Modeling Bibliographic Data in Library and DH Contexts," in which she will present the results of a survey of data use in DH projects examining the projects' use and presentation of library data (including bibliographic data and data presented in digital collections platforms). This survey will focus on

- whether each project includes bibliographic references;
- how much granular detail any bibliographic references include,
- whether or not such data is presented in the format of a specific model, and
- the presence or absence of links to specific copies, whether in library or digital archive catalogs (such as HathiTrust, the ESTC, paywalled collections (ECCO, EEBO, etc.), or websites like Google Books and Project Gutenberg).

Gathering this data will allow Morgan to look for common priorities in the project creators' use of library data; and to identify some of the assumptions that digital humanities has made about libraries and the bibliographic information that they produce. She will use the DH project survey data as the basis for a comparison with bibliographic ontologies (such as FRBRoo, BIBFRAME, Schema.org, and BIBO), and literature on challenges and best practices in bibliographic data modeling. Morgan will look for intersections, missed connections, and opportunities between digital humanities and cataloging work; and will consider how assumptions made in digital humanities elide the complexity and ongoing negotiations in the production of bibliographic data. In these projects, what drives the decisions to model (or not model) bibliographic data? How do the priorities of DH practitioners differ from those of library-based data creators? The increasing development of off-the-shelf tools, and the gradual growth of infrastructure for learning digital humanities skills and accessing data means that in many ways, it is possible for DHers, both new and experienced, to do more than before. However, increased access to materials and resources does not mean that the efforts of DH and library data communities will automatically complement each other. In FRBR, Before & After, Karen Coyle observes that library personnel working with data "have made

little change in our approach to subject analysis in the last half-century, possibly because there isn't a clear direction for improving this aspect of our work." This paper will argue that similar ambiguities exist around the use of bibliographic data in DH projects, and that the apparent common ground of bibliographic data use is more complex than it appears.