

Rushdie's Computers: Born-Digital Archives and Humanities Scholarship

Erika Leigh Farr (elfarr@emory.edu)

Emory University

As questions and concerns about digital preservation and sustainability become increasingly audible in the spheres of digital humanities and humanities computing, the necessity to build strong ties between digital humanities and digital libraries only intensifies. Howard Besser, in his essay, *The Past, Present, and Future of Digital Libraries*, underscores the connections between these two fields. He explains digital libraries not only “provide access to original source material, contextualization, and commentaries, but they also provide a set of additional resources and service”.¹ Besser then delineates some of these resources and services, including digital collections of traditional print materials, lexical analysis, and increased accessibility. In addition to these important contributions to humanities research, I would like to highlight the emerging role libraries play in processing, preserving, maintaining, and providing access to important archives that house born-digital content. This role will not only contribute to humanities research in decades to come, but will also impact how research is performed by directing what content is made available and how researchers may access it.

In this paper, I will examine the Emory University Libraries’ acquisition and processing of a singular personal archive as a case study to explore the methods and practices of handling born-digital archival materials and the implications such methodologies and their outcomes may have on humanities research.

Emory University’s acquisition of Salman Rushdie’s personal archive represents an important addition to the Manuscript and Rare Books Library, and contributes significantly to the University’s digital library resources and research. Rushdie’s rich personal archive includes traditional manuscript materials such as journals, personal correspondence, and notebooks, as well as less traditional archival materials, namely a series of personal computers that cover a significant span in his personal and literary life. This digital archive includes five computers, one early Macintosh desktop and four Macintosh laptops, including both obsolete and current models. While MARBL has previously acquired collections containing some digital materials, Rushdie’s computers represent the first significant,

sizeable digital component to the University’s extensive holdings of rare and unique materials.

Such an acquisition requires archivists to engage with technologists to ensure that the library can most effectively serve current and future researchers and scholars. The curation of such an archive raises important questions about how libraries should process, index, and present these materials while simultaneously addressing preservation and authenticity concerns. Such questions include: What is the research value of such an archive? How important is the physical artifact? Do researchers need access to exact systems emulation? Is providing search and browse access to the data sufficient or will researchers be interested in Rushdie’s original directory structure? Once data is migrated from the original environments, do we continue to maintain those outdated systems? How do archives sustain both master and access instances of born-digital archives?

As digital librarians and archivists at Emory begin processing the born-digital components of this important archive, they must keep these questions, and the host of secondary concerns circulating around them, in the foreground of workflow and process discussions. In this paper, I will argue for the importance of balancing the urgent needs of data and system stabilization with the more long-term challenges of considering the ideal outcomes and products of processing and providing access to a rare and unique born-digital archive. This talk will track the early stages of processing the physical and digital materials comprising Rushdie’s digital archive, outline approaches to handling the more complex processing requirements, discuss proposed approaches to presenting the archive to both local and distant researchers, and generalize observations drawn from the experiences with this born-digital archive to broader implications within digital libraries, digital curation, and humanities computing.

With the acquisition of any archive, a research library takes on multiple responsibilities to preserve, index, and provide access to the rare and unique materials. Libraries—digital, brick and mortar, or otherwise—must “incorporate the component of stewardship over a collection.”² Such stewardship carries with it important responsibilities, especially in cases of archival materials. Thus, with the arrival of the first shipment of Rushdie’s digital archive, consisting of three out of the five computers, our library and archive staff faced both immediate preservation demands and distant challenges for archival curation. We elected to produce a workflow that is both staged and modular, which I will only summarize here. The first task is to provide a secure and stable environment for the machines themselves. As our archives had not previously included technological artifacts, this first step required some intensive space and environment analysis. Once we stabilized the physical objects, this born digital archive next challenged staff with

questions of data recovery, data preservation, and data duplication. Such challenges prompted us to develop a partitioned data architecture that duplicates and preserves all master data. The original is preserved, a master duplicate is generated and stored darkly, a duplicate collection of the master database is housed in a secure repository for in-house processing and staging, and, finally, a fully-processed instance of the data is made available through a production database. Such an approach provides for preservation of original artifacts and master data, while ensuring a level of security for data while it is being processed for embargoed material.

In addition to preservation and security, authenticity of the archived data is of particular importance to archivists, digital librarians, and humanities researchers. Graham Barwell's discussions about originality and authenticity within the fields of textual studies and electronic textuality resonate with archives such as Rushdie's born-digital materials.³ How can we most authentically represent the digital archives included in the Rushdie collection? Is the data the only component of real research value or is the context that holds this data, the paratextual elements, if you will, of equal importance to researchers? I will explore these questions and provide illustrations from our processing of the Rushdie archive to offer some preliminary insights.

-
1. From Besser's essay, "The Past, Present, and Future of Digital Libraries" in *A Companion to Digital Humanities*, 2004, p. 557.
 2. Besser, p. 559.
 3. Graham Barwell, "Original, Authentic, Copy: Conceptual Issues in Digital Texts" in *Literary & Linguistic Computing* 20.4 (1995) see pages 416, 418, and 419.