

**Forensically Reconstructing Biomedical Equipment Technicians' Information  
Practices Using PDF Metadata: Toward a Conceptual Model of Layered  
Temporality**

James A. Hodges

Bullard Postdoctoral Research Fellow

University of Texas at Austin

School of Information

james.hodges@utexas.edu

**Abstract**

This study examines the documents circulated among biomedical equipment repair technicians in order to build a conceptual model that accounts for multi-layered temporality in technical healthcare professional communities. A bibliographic inquiry informed by digital forensics is employed to quantify and model the overlapping temporal, format-related, and annotation characteristics present in a corpus of device repair manual files crowdsourced in the course of collaborative efforts between volunteer archivists and professional technicians during spring 2020. The findings demonstrate that device repair manuals often travel through multiple steps of mediation between the time of their initial publication and their eventual usage by technicians, including the hand annotation

and scanning of paper documents, encoding as PDFs, porting forward into newer versions of PDF, and the addition (or removal) of new metadata over time.

By developing a conceptual model that accounts for multiple stages of such intervention, this project contributes to the study of working professionals' personal information practices and enables a more accurate assessment of the information needs associated with maintaining medical equipment—especially those related to working with devices and documents from a wide variety of temporal origins.

The corpus of documents used in this study originates within iFixit.com's Medical Device Repair collection, a crowdsourced trove of more than 10,000 device repair manuals contributed by working technicians in response to the strain placed on their colleagues and institutions due to the 2020 COVID-19 pandemic. The study focuses in particular on the Respiratory Analyzer subcategory of documents, which aid in the maintenance of equipment central to the care of COVID-19 patients experiencing respiratory symptoms. The 40 documents associated with Respiratory Analyzers in iFixit's collection are examined and quantified in terms of their original publication date, the apparent status of their original paper copies, the version of PDF used to encode them, and any additional metadata present in Extensible Metadata Platform (XMP) and/or Document Information Dictionary formats. After examining these bibliographic characteristics, the study turns toward developing a conceptual model that accounts for circulation of documents among multiple technicians, as well as alteration of documents during the course of their lifespans.

In conclusion, the author builds on these findings in order to consider areas where existing information practices, as reconstructed during the analysis, may

present opportunities for information specialists' support in terms of verifying authenticity, comparing multiple editions, and defense against intellectual property restrictions.

### **Preliminary findings**

iFixit.com's medical device repair manual collection was published in May 2020, with the stated aim of reducing strain on biomedical equipment technicians during the COVID-19 pandemic. Over 200 volunteer archivists aided in processing the collection, the largest publicly accessible collection of such materials available at the time of publication (Goode 2020). Today, a forensic bibliographic analysis of the materials can assist information professionals in better understanding the information practices of biomedical technicians, including the methods that they use to process, organize, and annotate documents critical in their professions.

In the top-level directory for Respiratory Analyzers, a series of 11 subfolders contain 40 PDF files documenting the operation and maintenance of devices crucial to monitoring patients' lung function when hospitalized with respiratory issues such as those often associated with COVID-19 (Respiratory Analyzers, N.D.). By examining the metadata in each file to reconstruct the personal information practices of biomed technicians, this study also develops a conceptual model capable of understanding the emergent information order produced during a global medical emergency.

One particularly rich example from this study's findings can exemplify the multi-layered temporality inherent in both biomedical technicians' professional activities and the documents that they collect in the course of such work. The iFixit

collection contains an instruction manual, in PDF format, for the Novametrix 7000 respiratory analyzer with its front page listing a publication date of July 20, 1989. The second scanned page of the Novametrix 7000 manual contains a list of handwritten notes enumerating maintenance costs and part numbers, showing that the manual is a living document of sorts, with marginalia used to enhance its usefulness in situ. These handwritten notes are also an example of the multi-layered temporality that is pervasive within iFixit's collection. In the course of a technician's career, they frequently collect and work on objects from a variety of times and places. The information order that emerges is unevenly distributed between rich and poor hospitals, old and new equipment, and technicians with larger or smaller collections of documents. As a result, any attempt at meeting the information needs of this population has to take into account its heterogeneity

An analysis of the document's XMP metadata provides additional detail on the multi-layered temporality of the Novametrix 7000 instruction manual. For example, despite its purported publication date in 1989, this digital file was actually generated in 2020 using version 1.7 of the PDF standard (initially introduced in 2018). Metadata also shows that the file was created using contemporary command line tools called PikePDF and Tesseract. Usage of command line tools demonstrates a degree of technical proficiency in the production and circulation of this particular document, which is not shared by all of the collection materials. In contrast, many other documents from the collection were produced using programs such as Apple Preview, the stock PDF viewer packaged with Macintosh operating systems. The presence of XMP metadata, furthermore, is only possible because the file was

generated after the XMP standard was adopted in 2001, with PDF version 1.4. This finding reveals another level of complexity and mediation, with a single object showing multiple levels of intervention. Overall, the findings show this particular document to have a long lifecycle. The constant interventions, in the forms of both hand annotation and digital encoding, also suggest that there is potential for outdated or maliciously altered files to populate technicians' personal records.

### **Annotated bibliography**

Fenlon, K., Senseney, M., Green, H., Bhattacharyya, S., Willis, C., & Downie, J. S.

(2014). Scholar-built collections: A study of user requirements for research in large-scale digital libraries. *Proceedings of the American Society for Information Science and Technology*, 51(1), 1–10. <https://doi.org/10.1002/meet.2014.14505101047>

This article provides a conceptual model for collection-building by academic researchers. My research builds on these findings by using forensic methods to understand how collections are developed in non-academic settings as well.

Finn, M. (2018). *Documenting Aftermath: Information Infrastructures in the Wake of Disasters*. MIT Press.

Finn's book introduces the concept of an emergency information order, or emergent social information system. This concept is useful to my study as I work to understand the ways that crisis in a professional community, such as COVID-19's effects on the biomedical technicians, lead to ad-hoc and complex systems of overlapping information exchange.

Gitelman, L. (2014). *Paper knowledge: Toward a media history of documents*. Duke University Press.

This book includes a chapter on the PDF format with useful historical and critical insight. As such it is useful in my efforts to understand the history of digital document production and circulation as it relates to biomedical equipment repair technicians.

Lee, H.L. (2000). What is a collection? *Journal of the American Society for Information Science*, 51(12), 1106–1113. [https://doi.org/10.1002/1097-4571\(2000\)9999:9999<::AID-ASI1018>3.0.CO;2-T](https://doi.org/10.1002/1097-4571(2000)9999:9999<::AID-ASI1018>3.0.CO;2-T)

By defining the notion of a collection more precisely, this article aids in my efforts to understand the organizational and thematic dimensions of bulk PDF resources such as iFixit's repair manuals.

Goode, L. (2020). *Right to Repair Groups Fire Shots at Medical Device Manufacturers* | *WIRED*. (n.d.). Retrieved September 14, 2020, from <https://www.wired.com/story/right-to-repair-medical-equipment-ifixit/>

Popular press article describing the iFixit biomedical device repair manual project.

*Respiratory Analyzer Repair*. (n.d.). iFixit. Retrieved September 30, 2020, from [https://www.ifixit.com/Device/Respiratory\\_Analyzer](https://www.ifixit.com/Device/Respiratory_Analyzer)

Main source of research data.

Shankar, K., Jeng, W., Thomer, A., Weber, N., & Yoon, A. (2020). Data curation as collective action during COVID-19. *Journal of the Association for Information Science and Technology*, <https://doi.org/10.1002/asi.24406>

Introduces the possibility of collaborative curation projects acting as useful response to COVID-19 pandemic. My research addresses a related but distinct issue, inasmuch as I address collective archival construction, rather than data curation as such.

Sullivan, S. J. (2006). An archival/records management perspective on PDF/A. *Records Management Journal*, 16(1), 51–56. <https://doi.org/10.1108/09565690610654783>

*Whitepaper: XMP Metadata support in PDFlib product.pdf*. (n.d.). Retrieved September 30, 2020, from

<https://origin2.cdn.componentsource.com/sites/default/files/resources/pdflib/545591/whitepaper-xmp-metadata-in-pdflib-products.pdf>

Both of these articles contain technical information on PDF document structure, XMP metadata standards, and PDF forensics. I build on this knowledge in my choice of a forensic/bibliographic methodology.

Ward, R., Wamsley, G., Schroeder, A., & Robins, D. B. (2000). Network organizational development in the public sector: A case study of the federal emergency management administration (FEMA). *Journal of the American Society for Information Science*, 51(11), 1018–1032. [https://doi.org/10.1002/1097-4571\(2000\)9999:9999<::AID-ASI1004>3.0.CO;2-5](https://doi.org/10.1002/1097-4571(2000)9999:9999<::AID-ASI1004>3.0.CO;2-5)

This article examines institutional policies related to emergency management and information systems. Methodologically speaking, it relies on official documents and interviews with agency personnel. I use this article in order to contrast my own approach, which focuses on response to emergency outside of established institutions, and uses forensic methodology in order to better understand the activities pursued by a geographically disparate community working largely without official attribution.