Preserving In-House Developed Software

Nicole Contaxis
National Digital Stewardship Resident
National Library of Medicine
Bethesda, MD
301-443-8000
nicole.contaxis@nih.gov

ABSTRACT

Software often plays a key role in the ways that institutions function, and some institutions, like the National Library of Medicine (NLM), have a history of developing software to serve their unique needs. Preserving this in-house developed software is the goal of the National Digital Stewardship Residency Project, "NLM-Developed Software as Cultural Heritage." This project will not only help ensure access to content created on this software, but, in the case of NLM, will also help document a long and often unrecognized intellectual history [1]. Although copyright concerns are largely avoided since the software was produced in-house, a variety of administrative obstacles still need to be addressed before the technical process of software preservation can begin. These obstacles include but are not limited to: (1) locating knowledge sources for software projects that are long defunct; (2) locating usable copies of software, either tangible or intangible, that may not have been properly documented or stored; and (3) tracing the history of projects that may have gone through several re-branding efforts or versioning's. This poster will address these issues as they have affected the current project at NLM and will demonstrate how a properly conducted inventory is necessary for contending with these obstacles and ensuring a reliable long-term software preservation strategy.

1. BACKGROUND

The National Library of Medicine has been developing software for internal and external use since the early 1960's when they began work on the Medical Literature Analysis and Retrieval System (MEDLARS). This computerized bibliographic system was meant to facilitate access to the library's bibliographic and serial records and to help compile the extensive indexes being produced at the time. With the creation of GRACE (Graphic Arts Composing Equipment), a custom phototype-setting machine, NLM was able to provide access to their data and print their indexes in record time. GRACE is now housed at the Smithsonian Institution [2].

In the 1970's with the creation and implementation of MEDLINE, NLM made their data available online, and years later, began to work on ways to offer full-text access through a variety of networks. In conjunction to their bibliographic systems, NLM experimented with a range of ways to satisfy the information needs of the health services community at large. Examples of such experiments include satellite communication to assist physicians in remote areas, the first internal library system, mobile computerized workstations to assist workers at toxic waste spills, user-friendly interface software for bibliographic access, and search and indexing features for GenBank and other influential databases [3].

The National Digital Stewardship Residency project aims to help preserve this history and call attention to the importance and impact of software development both within the library and beyond. The poster will represent a key part of this project as it illustrates how to deal with administrative issues using a thorough inventory before beginning the technical process of preservation.

2. OBJECTIVES

As this brief history of software development at NLM illustrates, software functions can vary wildly. Some projects will have user-interfaces while other will only perform computations. In order to make informed choices about how to preserve a software project, it is necessary to learn about that piece of software, how users interact or interacted with it, how it was developed, and how it changed over time. With the range of software projects from the history of NLM in mind, a delicate approach to the contexts of each project is necessary before committing to a preservation strategy for a particular piece of software. If, for example, a piece of software is meant to help users access and interact with information, it may be best to create an emulation in order to preserve it adequately. However, if a piece of software is important because of its computational uses, migration may be a better option. Furthermore, these types of decisions may be affected by the way a piece of software changes over time. Change occurs frequently for software projects, as versions and patches can affect the overall nature and experience of a project.

This poster intends to outline what information is necessary to inform these sorts of decisions, provide an example of how an inventory can be compiled, and illustrate how the peculiarities of software development can be accommodated in an archival setting. The inventory process necessitates close attention to the institution's history as well as to the ramifications of different preservation tools and techniques on the longevity of and meaning associated with a software project.

3. ACKNOWLEDGEMENTS

4. REFERNCES

[1] Library of Congress. 2015. National Digital Stewardship Residency. Webpage. http://www.digitalpreservation.gov/ndsr

[2] C.R. Dee, "The Development of the Medical Literature Analysis and Retrieval System (MEDLARS)," *Journal of the Medical Library Association.* 2007 Oct 95(4): 416-425.

[3] B. L. Humphreys, "Adjusting to progress: interactions between the National Library of Medicine and health science librarians, 1961-2001," *Journal of the Medical Library Association.* 2002 Jan 90(1): 4-20