

The National Security Archive

The George Washington University
Washington, DC

NDSR Project

The Digital Dissemination Challenge

Goal Summary

Carry out a diagnostic of and propose improvements to the National Security Archive's digital assets ingestion, management and dissemination processes.

Specific Goals / Objectives

The Resident would help us take a snapshot of all Archive activities that involve the capture, preservation and publication of digital assets: from the public reading room to analysts' offices; from the indexer/cataloger offices to the IT department, where heavy duty scanners and terabyte servers are operated; from the directors' offices to the Web masters' files, and those of our partners in other institutions which help us to post mega-collections on the Internet.

The resident will help us evaluate our processes in the fields mentioned above against the best practices suggested by established institutions and will present suggestions for improvements, reinforcement, and establishment of new procedures.

Timeframe & Deliverables

[Months 1-3] Conduct a review and analysis of current digital management operations and task flows, and identify solutions or improvements in these operations. This will involve the Resident working exclusively with the Information Systems Team (IST) [Director, Network administrator, Databases administrator and Web master] to get acquainted with the flow of digital content, operations and daily tasks as they stand, from scanning to storage and management of collections, applications and databases used in this process, to regular IT operations such as server management, data protection and security, as well as Web publication and content management systems.

[Months 4-6] Survey Archive staff on how they acquire, store, manage, and add value [metadata and organization] to documentary holdings; also, how they carry out electronic publication activities or on-site terminal serving of digital assets, as well as receive support from IST. To answer these questions, the Resident will approach the following staff members: the Reading Room and Front Desk staff; Indexers of the Production Team and the Torture Archive Project; Web Master and Blog coordinator; Project Analysts, Assistants and Interns involved in scanning, scraping or acquiring digital collections; senior staffers including the Director, Director of Research, Production head and key Board members and advisors involved in conceiving and implementing the goals and objectives of the Archive in general. Some of the issues to be explored at this stage will include:

- a) Technical: The way we acquire collections and information, and the way we disseminate our holdings; how to enhance the acquisition and dissemination processes by exploiting state-of-the-art methods or frameworks used in libraries or other similar leading organizations; how to use a Web-based approach and whether other platforms can be used to simplify and accelerate these processes.
- b) Organizational: How different staffers and teams interact; what regulations and procedures exist, or are needed for a smoother and more coherent, efficient and high impact operation.

[Months 7-8] The Resident with the help of Archive Directors will wrap up his or her work and produce a report with suggestions, key reinforcements, identification and modification of tasks, operations, work flows, guidelines and processes.

Resources Required

1 Mentor (Osorio), 1 Resident

Access to all staff: front desk-reading room-vault, scanning and IT team, indexers and cataloguers, web master, analysts and directors.

Contacts with organizations who have expertise in digital management, for example: Washington Research Libraries Consortium (WRLC), Benson Latin American Collection - University of Texas Libraries, U.S. Department of State Bureau of Information Resource Management, National Archives and Records Administration, Library of Congress.

Context

An Explosion in Holdings

By the turn of the millennium, the National Security Archive had established a reputation as the leading academic user of the Freedom of Information Act (FOIA), and a major independent repository of declassified U.S. government records on contemporary (post-World War II) history. At the time, we estimated that our collections – available to the public in our Reading Room – totaled 5 million pages. Of these, 40,000 documents (more than 200,000 pages) had also been published as richly curated microfiche collections available at educational institutions around the globe. We had also begun publishing concise compilations of selected documents in the form of “Electronic Briefing Books” (EBBs) on the new medium, the World Wide Web.

In the decade-plus since then, our paper holdings have expanded steadily. Today, researchers have access to more than 7 million pages of records acquired through FOIA, research trips to presidential libraries and other archives around the world, as well as through donations from authors, historians and journalists.

During the same period, our electronic records holdings have skyrocketed compared to what they used to be. We have targeted three specific portals for collection and broad dissemination in this format: our multiple issue web site containing more than 400 Briefing Books with analysis and supporting selections of U.S. records on varied national security and

foreign policy issues; the “Torture Archive” containing thousands of records pertaining to the Guantanamo Bay detention center and abuses in the Abu Ghraib prison; and the Nuclear Vault, which holds hundreds of key records needed to understand U.S. nuclear policy from the Cold war to the present. In all, these portals contain more than 10,000 annotated documents, along with analyses, which users can browse freely. In addition, researchers currently have access to a subscription product, the Digital National Security Archive (DNSA), which features more than 95,000 selected, cataloged and indexed documents (700,000 pages) that constitute important segments of the underlying record of U.S. foreign and military policy since the middle of the 20th Century. *The Washington Journalism Review* described these holdings collectively as “a state-of-the-art index to history.”

Since 2004, when we made our first major leap into the digital arena, we have multiplied the quantity of our digital holdings through:

- Scanning and ingestion of U.S. government responses to FOIA requests (100,000 documents);
- Capturing U.S. agency portals for permanent preservation purposes (700,000 documents)
- Ingesting large digital collections co-managed with international institutions (600,000 documents)
- Carrying out numerous scanning projects at archives around the country and the world (200,000 documents)

The Archive’s digital collections include records as varied and rich as the Paraguay Secret Police archives; Guatemalan Presidential Staff records; Truth Commission donations to Panama, Indonesia, Peru, Liberia and Rwanda; selections from former Soviet archives; and curated selections of the Kissinger papers (Memcons and Telcons); the Nuclear Vault; the Torture Archive (Guantanamo-Abu Ghraib); the Nixon Tapes; and Department of State records on Argentina, Chile, Guatemala and El Salvador.

In order to accomplish our second digital leap, the Archive faces a three-tiered challenge:

- 1) Bottlenecks in ingestion capacity: outlining ingestion processes; standardizing scanning parameters for different acquisition processes, etc.;
- 2) Management overload: selecting internal content management systems; standardizing digital formats and indexing; establishing robust container and backup processes; guidelines and procedures for precise accounting.
- 3) Bottlenecks in dissemination capacity: outlining dissemination means; acquiring and adapting content management publication tools.

*Required
Knowledge and
Skills for
Residents*

The resident will have a graduate degree in Library and/or Information Sciences, Information Technologies or equivalent from an accredited institution of higher education.

Additionally, the successful candidate will have the following:

General Knowledge

☐ digital conservation principles

Specialized Knowledge or Experience

☐ Application of digital preservation and/or digital asset management practices

☐ Descriptive metadata schema such as MARC, MODS or Dublin Core or their equivalent

Technical Experience

☐ Use of both Microsoft Windows and Apple Macintosh computers

☐ Use of office productivity software such as Microsoft Office and Adobe Acrobat

☐ Use of graphical design or imaging systems

*Preferred
Knowledge or
Experience*

- Familiarity with content management systems: Documentum, Laserfiche, Alchemy, Drupal, Greenstone, Dspace
 - Familiarity with scanning and digital acquisitions techniques
-