

Proceedings of the 7th Annual Federal Depository Library Conference

April 20 - 23, 1998

NTIS-GPO Imaging Pilot Project

Kristin M. Vajs, National Technical Information Service
Springfield, VA

The Vision

The American Technology Preeminence Act established procedures for Federal agencies to transfer unclassified information resulting from Federally funded research and development activities to the National Technical Information Service (NTIS). In the regulations implementing the act, NTIS proposed giving the depository library community free access to the agency's document image files as soon as it was technically feasible. NTIS offered to do this in the public interest, assuming all costs associated with preparing and disseminating the image files.

This vision was reiterated in greater detail in GPO's 1996 Study to Identify Measures Necessary for a Successful Transition to a More Electronic Federal Depository Library Program. In a document submitted to the working group, NTIS offered to provide depository libraries with access to its collection of documents in electronic format at no charge, as often as needed, and without time limitations. Depository libraries were only asked not to release the document image files outside the library or to allow the files to be used for commercial purposes. No restrictions were placed on the use or redissemination of printed documents. The proposal also suggested that a group of 20 depository libraries test the feasibility of the concept.

The Vision Becomes Possible

In the fall of 1996 important developments occurred which enabled us to begin implementing the vision.

1. Electronic Ordering via the Web

NTIS put into place an online ordering capability called OrderNow Online. This Web-based service provides a mechanism for placing and transmitting document orders electronically, 24 hours a day, 7 days a week. OrderNow Online covers documents received by NTIS during the last 90 days.

2. Image File Availability

Another factor which increased the viability of the project was a substantial increase in the number of documents available in electronic format. In September 1996 the Department of Energy joined the Defense Technical Information Center (DTIC) in transmitting documents to NTIS in image format. At this time, NTIS also began scanning documents in the collection for which orders had been received.

3. A Test Site

Linda Kennedy offered the depository library at the University of California (UC) at Davis as the first test site. The participation of the UC-Davis Library was crucial because it allowed us to demonstrate the feasibility of the vision prior to orchestrating the extensive interagency support necessary for expanded testing.

The Pre-Pilot Project

Several months of conference calls between Linda's team at the University of California at Davis and the NTIS team followed. Meticulous planning covering both technical and policy issues was necessary before the pre-pilot project could begin. We sought solutions to such problems as how NTIS would indicate which documents were available in image format, how UC-Davis would make the documents available to users, and how an order-taking system that was prepared to accept either credit card or deposit account numbers could be successfully adapted to this project.

At NTIS, staff in seven different offices in four directorates were involved in either advising or working on the UC-Davis project. One of the biggest challenges we faced was trying to link stand-alone automated processes into an integrated system. Many of the pieces of what we were trying to do were there--we could take an order electronically, we could create an image file, but automated links between these processes did not exist. Human intervention was ultimately necessary to run the pre-pilot project. Fortunately, ongoing development work at NTIS will transform this situation.

Once the feasibility of the endeavor was demonstrated, many people at GPO and NTIS worked to get administrative support the project needed. NTIS wanted the pilot project to be operated under the sponsorship of the Federal Depository Library Program. On the road to getting the interagency agreement governing the pilot project signed, there was consultation with the NTIS Advisory Board, the Depository Library Council, the Joint Committee on Printing, and discussions at the highest levels at GPO and NTIS.

During the past year we have done more than send documents to UC-Davis: we developed the infrastructure, technical and administrative, which now allows us to broaden participation in the project.

Time for Another Quantum Leap

There are a number of technological and work process developments taking place at NTIS which will substantially benefit the pilot project.

1. Increase in number of documents available in image format

The range of documents included in the project was dramatically expanded when NTIS began scanning virtually all incoming documents in October 1997. Documents are scanned immediately after it is determined that they are unique to the collection. Cataloging and indexing follow directly afterwards. This new work process greatly increases the number of documents available through the pilot project. We estimate that approximately 40,000 Government documents are currently available in image format, and this number is being increased by nearly 3,000 each month. All incoming documents in the "PB" series are now scanned, except in the few cases where full color graphs, charts or maps preclude such treatment. This scanning also improves the timing of NTIS document delivery functions, because it assures that documents are available in any format, including image files, at the point when they are announced through commercial versions of the NTIS Database or through our Web site.

Our capacity to deal with incoming documents in electronic form is also being enhanced through the development of an electronic document processing workstation. We are receiving an increasing number of requests from agencies that are interested in sending their documents to us electronically. We have systems currently in place to deal efficiently with large scale electronic transmission--collections from DTIC and DOE--but our capacity to deal with a few documents coming from many places has been limited. Our new workstations will allow us to retrieve, view and catalog documents and then send them on to a conversion workstation connected to the document storage system. Documents received in word processed or PDF formats will be converted to the TIFF format before storage as part of the NTIS electronic archives.

2. Fully Automated Delivery of Image Files

The linking of our automated ordering service with the automated generation of an image file is also imminent. In May, NTIS will launch a new service, which is driving the development of tools which will also be useful for the depository library project. TDP/MIS (Technical Data Package/Material Information System) is a Web-based document ordering and delivery system. Government contractors and manufacturers will be able to search for Department of Defense solicitations of interest and then view online the entire technical package requirements for the solicitation. The technical package consists of industrial standards and military specifications that products provided by bidders must meet. The TDP/MIS ordering component includes a shopping cart where the user places documents while browsing the system. At checkout time, the user pays for the documents with a credit card, electronic money or a deposit account. The documents can then be downloaded by the user immediately.

Generalizing capabilities developed for the TDP/MIS project, in the near future we hope the pilot project will work as follows:

- a. The user will identify a document using the Technical Reports File on the NTIS Web site and submit an order electronically. The Technical Reports File covers NTIS document receipts over the past ten years. (By using the

- Technical Reports File rather than OrderNow, the number of documents libraries have access to will increase greatly.)
- b. The ordering library will receive an e-mail message indicating that the order has been received.
 - c. The ordering library will receive an e-mail message, only hours after the first message, saying that the document is ready to be picked up on the NTIS server.
 - d. Participating libraries will decide whether to make documents available through a directory; transmit them directly to the requester on their local area network; or print the document.

NTIS plans to fully automate the request fulfillment system that will allow requests to be submitted and filled 24 hours a day, 7 days a week. Automation will expand the hours the service is available and improve response time, particularly for libraries in other time zones. NTIS staff time will only be necessary in setting up new participants or if system problems arise.

5. Safeguarding the Electronic Archives

NTIS is also taking steps to safeguard its electronic archives. A tape backup is prepared for each document that we scan. Within the next 6 weeks we will have an offsite storage site for the backup copy. TIFF images are currently the de facto industry standard. TIFF was also the format choice made by the CENDI agencies, an interagency information planning organization which includes DOE, DTIC, NASA, and NTIS. However, NTIS will also have the capacity to convert documents to other formats should that prove necessary.

6. Project Communication

Over the course of the pre-pilot we have used GOVDOC-L as the primary way of keeping the Government documents community informed about our progress. NTIS is developing a Web page geared to the interests of libraries and information professionals. The page will feature up-to-date information on the pilot project as well as offering a virtual tour of document processing and ordering operations. Additionally, NTIS will host a listserv for pilot project participants to facilitate the sharing of questions and information. Facilitating communication among project participants will make it easier for selected libraries to get started and easier to solve any problems that may arise.

Future participants will have the option of picking and choosing among several ways of running the project identified by the early participants or of developing their own paths. Librarians will make those decisions based on what's best for their institution, the level of technical expertise available and the amount of time available to devote to the project. Later participants will benefit from having a peer group to assist them in getting started, as well as help from NTIS and GPO.

Conclusion

This project presents an alternative model for providing access to Government information via the Internet. If resources are unlimited or document collections relatively small, then documents can be maintained on Web sites indefinitely. NTIS has proposed a cost effective way of providing the depository library community with access to a large Web catalog of documents, the full electronic text of which can be provided quickly and in a standard format, on request.

The NTIS-GPO Imaging Pilot Project: Q&A	
What types of documents are covered by the project?	Documents produced by the U.S. Federal Government which are available on the NTIS document imaging system are included in the pilot project. Most of the documents available were issued from 1996 to date. Materials which are copyrighted or which require NTIS to pay a royalty to the document creator are not included as part of the project. NTIS-produced CD-ROMs, datafiles, audio-visual products and databases are not covered by the program.
How many documents will be available to pilot participants?	Currently, we estimate that we have about 40,000 documents of the type described above on the NTIS document imaging system. This number will grow at the rate of approximately 3,000 a month.
How will institutions be selected for participation in the pilot project?	GPO will select the institutions that will participate in the pilot project.
Are there any costs for participating?	NTIS will assume all costs associated with handling the requests and creating the image files. Participating institutions will need to absorb or charge the user for the cost of printing documents.
What is the timeframe for adding libraries to	Three or four libraries will be added in

<p>the pilot project?</p>	<p>a June-July 1998 time frame. Additional libraries will be added when the request fulfillment process has been automated.</p>
<p>What are the characteristics of the image files?</p>	<p>The average NTIS document--about 100 pages, TIFF images in a Postscript wrapper--is roughly five megabytes in size. These documents will take time to download and space to store. Viewing online may be slow, depending on the type of equipment a library has. Ghostscript or Ghostview is a public domain TIFF viewer that will run on most PCs.</p>
<p>In general terms, how will the project work?</p>	<ol style="list-style-type: none"> 7. The user will identify a document using the Technical Reports File on the NTIS Web site and submit an order electronically. The Technical Reports File covers NTIS document receipts over the past ten years. 8. The ordering library will receive an e-mail message indicating that the order has been received. 9. The ordering library will receive an e-mail message, only hours after the first message, saying that the document is ready to be picked up on the NTIS server. 10. Participating libraries will decide whether to make documents available through a directory; transmit them directly to the requester on their local area network; or print the document.
<p>How may a library or a library's clients use the documents received?</p>	<p>Within the library's community, patrons may view, print, or download</p>

	<p>the images. The documents may be made accessible from an institution's intranet or institutional network. The documents may not be posted on servers accessible to persons outside the library's user community.</p>
<p>How will the pilot project be evaluated?</p>	<p>GPO and NTIS, with input from participating libraries, will develop an evaluation plan. NTIS may also elect to study the impact of project participation on sales in a given geographic area.</p>
<p>Are there restrictions on the number of documents that can be ordered?</p>	<p>Participants are asked to limit their orders to 5 documents per week until the request fulfillment process has been fully automated.</p>
<p>What will happen at the end of the pilot project?</p>	<p>NTIS and GPO are optimistic that the project will be successful and are hopeful that a successful outcome would result in the continuation and expansion of the project. The interagency agreement between NTIS and GPO does not have an ending date, meaning that the project will continue as long as the two parties to the agreement are willing to continue it.</p>