



From Prints to Bits

Digitizing NIST's Archival Photograph Collections

Katelynd Bucher and Andrea Medina-Smith, National Institute of Standards and Technology

The National Institute of Standards and Technology (NIST), founded in 1901 as the National Bureau of Standards (NBS), is one of the nation's oldest physical science research and metrology laboratories. It has been instrumental in the development and standardization of many measurements and technologies Americans are familiar with today, such as standard weights and measures, radio navigation, and computer security.

The Information Services Office (ISO) houses the NIST Research Library; the NIST Museum; and archival collections of photographs, oral histories, agency publications, and other materials that record NIST's history. Since 2008, ISO has undertaken several digitization projects to increase the visibility and access to these noteworthy collections.

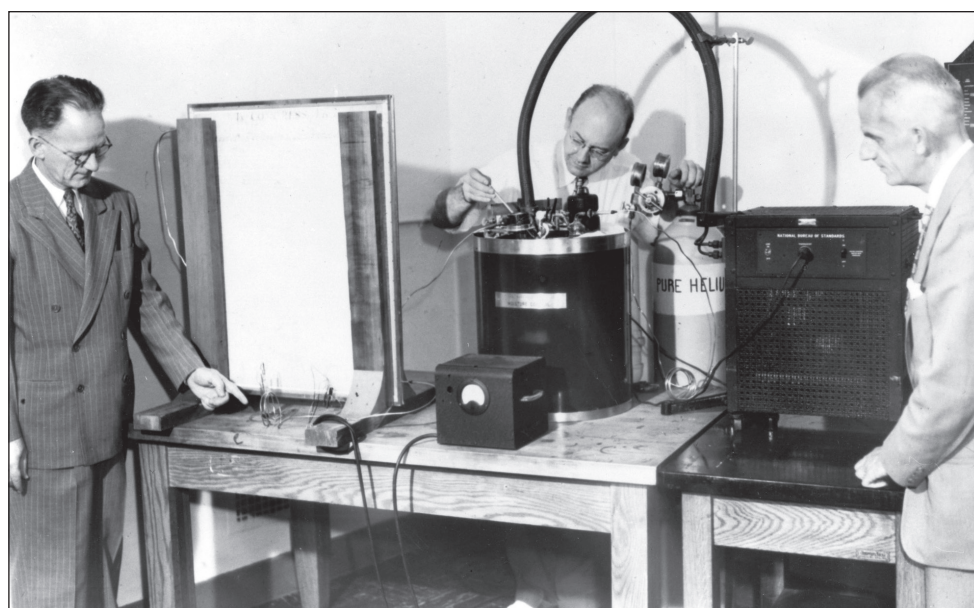
NIST Archival Photograph Collections

The photographs in the NIST ISO archival photograph collections visually document the research NIST has performed since 1901. There are more than 150 archival photo collections, and that number continues to grow. The photos document NIST's history; notable NIST leaders and researchers; and a variety of

research projects, tools, and technological advancements.

In 2012, ISO began digitizing the archival photograph collections and making them available on the NIST Digital Archives (NDA) at <http://nistdigitalarchives.contentdm.oclc.org>. In the last year, more than forty

Above photo: Between 1920 and 1930, the number of cars registered in the United States leaped from 9 to 26.5 million. NBS research on the automobile and the airplane began as an effort to conserve the nation's supply of gasoline and oil in hopes that better knowledge of fuels, ignition, lubrication, and carburation would assist in lowering the gasoline consumption of automobiles. *Courtesy of the National Institute of Standards and Technology.*



In 1939, after an investigation undertaken at the request of the Librarian of Congress, NBS recommended that documents be placed in specially constructed enclosures, that the air in the enclosures be replaced with a chemically inert gas, and that the enclosures be sealed. New enclosures were developed by NBS and the Libbey-Owens-Ford Glass Company. In 1951, the enclosures were completed and the Declaration of Independence and the Constitution of the United States were sealed within them in a ceremony with President Harry Truman. *Courtesy of the National Institute of Standards and Technology.*

archival photograph collections have been posted. The collections are organized by research topic and mainly consist of 8"x10" prints and accompanying negatives created both by official NIST photographers and contracted photographers to document NIST research and events throughout the last century. Until this point, the photo collections were mainly used by NIST research librarians, the NIST Public Affairs Office, and by members of the NIST Standards Alumni Association.

Nearly all the physical collections have one or two photos with a large amount of metadata attached to them, while the rest of the photos have at least a title or caption. The metadata consists both of typed information taken from internal publications and handwritten notations with photograph numbers and references to further resources. The dates, information identifying pictured individuals, and photograph titles, as well as biographic or historical notes, have been arranged by previous NIST librarians who gathered all the metadata and arranged it with the photographs in collections by subject in the early 2000s.

The archival photo collections are catalogued in the NIST Research Library's online catalog at the collection level. The collections' arrangement, brief descriptions included with the photos, and collection-level cataloging over the past decade have laid the valuable groundwork for the current digitization project.

Project Goals and Achievements

The overarching goal of the digitization project is to increase the visibility of NIST's work by making its archival photo collections more easily accessible to the public and the NIST community. To accomplish this, ISO is adding the digitized photograph collections to the NDA as the digitization and metadata creation processes are completed. ISO is also in the process of creating finding aids for each collection to provide additional information.

In 2013, ISO completed its goal of digitizing the most frequently requested archival photograph collections. The goal for 2014 is to digitize the remainder of the archival photograph collections, make them available on the NDA, and more widely publicize the photo collections.

Step by Step

The project team consisted of two metadata librarians and three support staff (an information specialist, editorial assistant, and a digital composition specialist). In addition, two senior librarians with strong cataloging and writing skills as well as in-depth knowledge of NIST's history and research provide guidance and review the finding aids and collection descriptions.

The photographs pass through several steps within the digitization process. First, the photographs are scanned at 600 dots per inch (dpi) in TIFF format. Each scanned photograph is then checked for image quality; we ensure it is straight and complete and without excessive margins and white space. The image is then adjusted or rescanned as necessary. Second, a metadata librarian uses the metadata attached to the print copies of the photos and additional historical resources to create metadata for each photo at the item level using Adobe Bridge's¹ metadata functions. These descriptions are added in two passes: first, the metadata librarian adds a metadata template containing collection-level fields,

Continued on page 25 >>

Only **GAYLORD** can offer you

GUARANTEED IN STOCK

YOU WANT IT? WE'VE GOT IT.

(Or we'll take **20% OFF** the cost of that item!)

Products designated as guaranteed in stock ship same day* with **no expediting fees** and **no minimums**.



For more information visit
www.Gaylord.com/GuaranteedInStock

Gaylord

Your Trusted Source®

CALL: 1-800-448-6160 | FAX: 1-800-272-3412 | EMAIL: CUSTOMERSERVICE@GAYLORD.COM

*Guaranteed in-stock items ship same day if order is placed before 2:00 pm EST, Monday-Friday. Some restrictions apply.



From Print to Bits continued from page 5

including the creator, source, copyright information, attribution statement, and keywords, which are taken from the Library of Congress Subject Headings database; then, item-level metadata is added, including a title and, if available, date created, identifier, and description.

Third, the metadata librarian then uses Adobe Photoshop's image processor to create JPEG copies of the photographs. The metadata is then embedded into the photos, ensuring that all available contextual material will remain with the photographs as users discover them, whether on the NDA or elsewhere. The metadata will remain within the photos no matter where they are located. The metadata then goes through another quality control process to ensure accuracy, and the photos are uploaded by collection to the NDA, where they are available to the public.

The last step of the process is to create a finding aid for each collection. The collections are small and described at the item level within CONTENTdm, and at a basic collection level in the library's online catalog. The finding aids add contextual information in an easily downloadable format (PDF), but do not preclude using the images on their own. Because the work involved in writing the finding aid is a separate workflow, these are not necessarily completed in the same timeframe as the digitization and upload processes.

Documentation of the project's processes (including scripts and XSLT transformations in a zipped file) can be found at the Information Services Office's website, the NIST Virtual Library (NVL), at <http://nvlpubs.nist.gov/nistpubs/documentation/ProcessingPhotoCollectionsForPublic.zip>.

Tools, Tips, and Tricks Learned Along the Way

- Automate where possible. Letting a computer perform bulk transformations and complex data manipulation is much more resource effective than doing each by hand.
- Ask for help! Listservs like Code4Lib and XML4Lib are great resources, even if you are inexperienced at scripting and coding.



National Geographic and NBS jointly sponsored an expedition to the Kazak region of Russia (then U.S.S.R.) to observe the solar eclipse of June 1936. *Courtesy of the National Institute of Standards and Technology.*

- Start with small collections that already have some written descriptive information. If you don't have any, think about creating a small collection and crowdsourcing to obtain rich descriptions.
- Make metadata templates in Adobe Bridge before you begin the process of describing your images. It makes more sense to apply the collection-level information to all the images at once and then add more fields rather than repeatedly cutting and pasting.
- Use the Tools → Photoshop → Image Processor button in Adobe Bridge to quickly batch transform images from archival TIFs to access JPEGs.
- The open source and free product ExifTool uses either the command line or a graphical user interface to extract, add, or change metadata embedded in nearly any type of file format. It was designed for use with EXIF metadata fields, but works just as easily with XMP, DC, PDF, and Photoshop fields. ISO uses this tool to pull XML records about each image file.
- Once XML files are produced they can be transformed into any format needed for upload to a digital repository. In this case, ISO uses CONTENTdm so the XML is transformed into a tab-delimited text file.

Next Steps

ISO's photo digitization project is ongoing. The team started with the most-requested collections in the first stage of the project,

and the second stage will be comprised of collections that are used less frequently. Digitizing the less frequently requested photographs should increase their visibility and use. Simultaneously, ISO will be enriching the metadata and the finding aids already on the NDA.

To increase awareness of the digitized photograph collections within NIST, ISO periodically announces their availability in the news section of the internal NIST website and the NVL. ISO also highlights specific photos from the collections on the homepage of the external NVL (www.nist.gov/nvl). Beginning in 2014, ISO will be adding selected collections to Wikimedia Commons. ISO will monitor and report on the use of the collections, which are essential aspects of determining and demonstrating the value of the investment in the project.

The work continues, but it has been a rewarding project for NIST and ISO. We hope digitization will give more researchers and the public better access to NIST's inspiring and impactful work. ■

Notes

¹ Certain commercial products are identified in this article in order to specify the procedures adequately. Such identification is not intended to imply recommendation or endorsement by the National Institute of Standards and Technology, nor is it intended to imply that the products identified are necessarily the best available for the purpose.