Short Communication

DICOMs, Missiles, and Metadata: The U.S. Nuclear Weapons Program Leverages a Medical Standard

Laura McGuiness

Metadata and AI Applications Librarian
National Security Research Center Mission Support
Los Alamos National Laboratory
Los Alamos, New Mexico, USA
lauraleemcguiness@gmail.com
https://orcid.org/0009-0002-5232-2557

Abstract

Digital Imaging and Communications in Medicine (DICOM) images, although most commonly used in medical settings, have been widely adopted by the United States Department of Energy (DOE) for capturing images of the internal components in a nuclear weapon. DICOMs, a lesser-known file format that combine nested metadata structures with complex image data—including multiple planes, frames, and high resolution—require the creation of access copies to support usability within the DOE. Used primarily for ensuring the safety, security, and reliability of the U.S. nuclear stockpile, the Los Alamos National Laboratory (LANL)'s DICOM images and corresponding image metadata must be accessible to scientists and researchers via our institutional centralized databases. This paper describes the author's creation of a Python script that converts DICOM images into accessible, archive-friendly TIFF files while preserving key image data and metadata.

Introduction

The Los Alamos National Laboratory (LANL) is a United States national laboratory responsible for ensuring the safety, security, and effectiveness of the US nuclear stockpile. LANL's National Security Research Center (NSRC) is the Lab's classified library that houses tens of millions of materials related to the development, testing, and production of nuclear weapons. Tracing its lineage back to the technical library formed by J. Robert Oppenheimer in 1943 as part of the Manhattan Project, the NSRC is one of the largest scientific and technical libraries in the federal government (Los Alamos National Laboratory, n.d.). Materials span from the 1940s to the present, ranging from paper documents and microform to more unusual formats like radiographs.

LANL's collection of radiographs mostly stems from the Pantex Plant, which uses "X-rays, lasers, dye penetrants, eddy currents, coordinate measuring machines and other technologies" to perform non-destructive testing that ensures "the safety and quality of the nation's nuclear weapons stockpile" (Pantex ,n.d.-b, "Microfocus Radiography", para. 1.). COLOSSIS, the world's only X-ray inspection system, is designed to look inside the core of a nuclear weapon, allowing scientists to analyze the images for any problems or anomalies (Lawrence Livermore National Laboratory, 2022; Pantex, n.d.-a).

While our collection houses a large number of physical radiographs, digital files have posed significant problems due to their unique file format. The Digital Imaging and Communications in Medicine (DICOM) file type is an international standard for transmitting, storing, retrieving, and displaying medical information. DICOM format allows for both image and metadata to be packaged into a single file. Due to the format's widespread use in the medical industry, the DICOM standard relies heavily on metadata elements aligned with interpreting medical images.

Additionally, a single DICOM image can be both multi-frame and have multiple overlay planes, representing different layers of a scan (DICOM Standards Committee, 2025). LANL researchers working with DICOM files create large three-dimensional images that can be viewed from different angles. This capability is used to detect structural variations and can ensure inspections of the stockpile generate high fidelity surveillance data for scientists across the National Security Enterprise (Lawrence Livermore National Laboratory, 2022).

Context/Problem

The DICOM images described in this study were harvested from a secure shared drive and ingested into our institutional digital repository, National Security Data Solution (NSDS). Due to the limitations of searching within a shared drive, the images were difficult to find due to the limitations of searching across subdirectories with little searchable metadata and non-descriptive file names. The NSDS offered the ability to store preservation DICOM images in perpetuity, but access copies were not initially available for easy browsing. Users needed to download each DICOM file and subsequently view the image using a specialized DICOM viewer or image editing software (Graham et al., 2005, p. 1138). To render the files more accessible, I converted each DICOM into TIFF format, preserving the same high level of resolution, fidelity,

and 3D viewing capability of the original DICOM. This afforded NSDS users the ability to view radiographs within the NSDS before choosing to download the full DICOM image.

Ensuring our users' ability to leverage metadata in their search functionality was critical to promoting the accessibility and value of these images. While users could view individual DICOM file metadata within photo editing software, I felt it was necessary to generate a metadata file that would provide additional search criteria within NSDS overall. Components of DICOM's standardized metadata structure are specific to medicine (e.g. patient/study information); however, metadata for each image also includes series/acquisition, image, and equipment information that can capture the unique historical, technical, and scientific context needed by LANL researchers who rely on this data for decision-making and recommendations regarding the U.S. nuclear stockpile.

Tools and Approach

To support this project, this author created a standalone Python script that could run without a graphical user interface (GUI) or any external software dependencies. The script relied on a combination of the following Python libraries: pydicom, Pillow (PIL), NumPy, and os. Pydicom read and parsed the DICOM files and enabled direct access to DICOM tags like PixelData, SamplesPerPixel, etc. Pillow (PIL) was used to convert NumPy arrays into TIFF images and supported multiple image modes as described in Figure 1. NumPy handled pixel data arrays and rescaled pixel intensity as needed.

The os library allowed the script to recursively walk through a directory of DICOM files and then store TIFFs as output files in a mirrored folder structure named after the parent of each DICOM folder subdirectory. A custom substring was extracted from the DICOM folder name and used in the output filenames for accessibility and connectivity to the preservation DICOM files. This ensured consistent traceability between source and converted files.

The script dynamically identified technical characteristics of each DICOM file using BitsAllocated, BitsStored, SamplesPerPixel, and PhotometricInterpretation DICOM tags. Values of SamplesPerPixel and interpretations such as MONOCHROME1, RGB, and YBR_FULL distinguished between grayscale and color images to allow for the appropriate TIFF output format. Depending on the properties detected, images were then converted as described in Figure 1.

Table 1. The left column is the logic applied during DICOM parsing. The middle column is the interpretation of the image type. Both right columns are the type of TIFF and the image mode used by the Python Imaging Library (Pillow).

Detected Property	Determination	TIFF Output	Python Image Mode
SamplesPerPixel = 1 and BitsAllocated ≤ 8	8-bit grayscale	Rescaled to 0-255	"L" (8-bit)
SamplesPerPixel = 1 and BitsAllocated > 8	16-bit grayscale	Rescaled to 0-65535	"I;16" (16 bit)
SamplesPerPixel = 3 and PhotometricInterpretation = RGB	8-bit color (RGB)	Channels preserved as-is	"RGB"
SamplesPerPixel = 3 and PhotometricInterpretation contains YBR	8-bit color (YBR to RGB)	Converted to RGB before saving	"RGB"
Format not matching known patterns	Unknown format	Defaulted to 8-bit grayscale	"L" fallback

Outcome and Challenges

Numerous challenges arose during the process of developing this scripted workflow. Most of these obstacles revolved around maintaining image quality during the conversion process. I encountered several problems detecting the original DICOM format and converting it into the appropriate TIFF counterpart. Any other issues resulted from the unusual organization of metadata within a DICOM file. All of these challenges were resolved after careful research into DICOM images and metadata and querying of tools like StackExchange and ChatGPT.

Post-conversion, a small portion of the TIFF files presented unanticipated behaviors. Some were saving as 8-bit when the source was 16-bit, others were generating all black images, and the script skipped or output errors when encountering unknown or nonstandard DICOM files. While reviewing outputted TIFFS, all 16-Bit DICOM files appeared low-contrast. This issue was resolved by rewriting the script to inspect the BitsAllocated metadata field in each DICOM file; if this field indicated it was over 8, the script set the correct image mode in the Python Pillow library and ensured the pixel array was set to 16 bits. A similar process was followed once it became obvious a subset of TIFFs were underexposed, appearing visually as blank or all black. The script rescaled the pixel values to the full 16-bit range by writing a function that rescaled the pixel array as needed. Finally, some DICOM images lacked tags like PhotometricInterpretation or assigned odd values to SamplesPerPixel. As a general rule, the script defaulted to 8-bit grayscale and logged these issues. This path was taken to ensure the converted images were preservable and usable even when perfect conversion wasn't possible.

Most metadata problems resulted from the hierarchical data representation unique to the DICOM format. Early in the conversion workflow, the script began struggling to process oversized metadata files, which can contain up to 5,000 distinct metadata fields (Microsoft Corporation, 2025). Nested sequences, which represent hierarchical data

within a DICOM file, are distinct from traditional metadata fields that exist as name-value pairs. Instead, they list structured data items that contain their own DICOM metadata fields. Each metadata attribute is also assigned a Value Representation (VR) that defines the data type and format (DICOM Standards Committee, 2025). Due to the potential of bloated metadata files, this author adjusted the script to exclude certain unwanted fields as well as output the files in separate documents.

Within the script, each DICOM file was examined for nested sequences by looking for a VR equal to "SQ" and then recursively traversing each sequence (if one existed). I then extracted all nested metadata elements and constructed the outputted metadata file into a human-readable format. To continue ensuring readability of the metadata files, I excluded PixelData, a metadata field that represents the binary pixel matrix of the image itself that was being retained already within the TIFF files. While ideally this metadata would have been kept, its presence in the TXT files slowed down processing and storage, negatively affecting the accessibility of the outputted metadata files.

In total, 1,243 DICOM images were converted into the equivalent number of TIFF and TXT files allowing for both modularity and readability. The script automatically saved each file's metadata in TXT files and placed them alongside the converted TIFFs in the matching subdirectory, ensuring ease of access during ingestion into the NSDS. Within the repository, individual records included 3 files each: the DICOM file complete with embedded metadata (kept for preservation purposes), the TIFF file (created as an access copy), and a plain text file containing all the extracted metadata.

The metadata captured during digitization adheres to the Digital Imaging and Communication in Non-Destructive Evaluation (DICONDE) metadata standard. Established in 1993, DICONDE standardizes all metadata fields and elements in a universal format, allowing for interoperability across many users and vendors (Brunnengraber, 2018). Because this standard was created for wide use within the medical field, Department of Energy radiographers have repurposed this schema for our use to ensure adherence to DICONDE. For example, digitizers captured pertinent metadata such as weapon system under Patient ID and part serial numbers under Patient Name. The use of DICONDE, while used unconventionally, is a pragmatic approach that supports accessibility and searchability. Ultimately, the TIFF files will contribute to greater interoperability of radiograph data amongst the United States Department of Energy.

Conclusion

After troubleshooting the aforementioned challenges, the Python script worked as expected and ultimately provided access to 1,243 DICOM images by providing TIFF images comparable in resolution. One of the final steps was to consult subject matter experts (SMEs) to evaluate whether any image quality, resolution, or fidelity of any kind was lost during conversion. The SMEs used side-by-side comparisons with the original DICOMs and assessed whether any loss in structural detail was apparent after the conversion process. In addition to equivalent image fidelity with greater levels of access, the extracted metadata files ensured access without needing specialized software. As of now, the DICOM files ensure access to NSDS users, allowing users to

gather data that in part ensures the safety, security, and reliability of the U.S. nuclear weapons stockpile.

Acknowledgements

Disclosure Statement

Funding

References

Brunnengraber, P. (2018, July 8). DICONDE in digital NDT taking hold. *Quality Magazine*. https://www.qualitymag.com/articles/94832-diconde-in-digital-ndt-taking-hold

DICOM Standards Committee (2025). *DICOM standard Part 5: Data structures and encoding* (PS 3.5) [PDF]. National Electrical Manufacturers Association. https://dicom.nema.org/medical/dicom/current/output/pdf/part05.pdf

Graham, R. N. J., Perriss, R. W., & Scarsbrook, A. F. (2005). DICOM demystified: A review of digital file formats and their use in radiological practice. *Clinical Radiology, 60*(11), 1133–1140. https://doi.org/10.1016/j.crad.2005.07.003

Lawrence Livermore National Laboratory (2022, April 5). *Scientists develop new CT scanner to image nuclear weapon components*. https://www.llnl.gov/article/34751/scientists-develop-new-ct-scanner-image-nuclear-weapon-components

Los Alamos National Laboratory (n.d.). *National Security Research Center*. https://nsrc.lanl.gov/

Microsoft Corporation (2025, March 31). *DICOM data transformation considerations*. Microsoft Learn. Retrieved August 1, 2025, from https://learn.microsoft.com/en-us/industry/healthcare/healthcare-data-solutions/dicom-data-transformation-considerations

Pantex (n.d.-a). *Colossis*. U.S. Department of Energy. https://pantex.energy.gov/mission/technology-development-and-transfer/colossis

Pantex (n.d.-b). *Microfocus radiography*. U.S. Department of Energy. https://pantex.energy.gov/mission/technology-development-and-transfer/microfocus-radiography