ISSN: 2572-3235

Burbridge. Int J Radiol Imaging Technol 2020, 6:065
DOI: 10.23937/2572-3235.1510065
Volume 6 | Issue 2
Open Access

International Journal of
Radiology and Imaging Technology

BRIEF REPORT

Dicom Image Anonymization and Transfer to Create a Diagnostic Radiology Teaching File

Brent Burbridge, MD, FRCPC

Medical Imaging, Royal University Hospital, University of Saskatchewan College of Medicine, Canada

***Corresponding author:** *Dr. Brent Burbridge, Medical Imaging, Royal University Hospital, University of Saskatchewan College of Medicine, 103 Hospital Drive, Saskatoon, SK, S7N 0W8, Canada, Tel: 1-306-655-2410*

Abstract

The creation of a digital, Diagnostic Radiology, teaching file system is reliant upon a process for selecting, anonymizing, and exporting Digital Image and Communications in Medicine (DICOM) images from a clinical Picture, Archive, and Communication System (PACS) to the teaching file.

A local version of the Medical Imaging Resource Center - Teaching File System (TFS) from the Radiology Society of North America (RSNA) was deployed to create teaching files. Philips, Intellispace PACS, was the source of the DICOM images.

The image capture, anonymization, and export processes to prepare images from PACS for the TFS were mediated by a web-based application, the Teaching File Transfer Tool (TFTT).

Technical strategies for protecting the confidentiality of patient information when building a Diagnostic Radiology digital teaching file are presented. The educational benefits of these technologies have a significant impact upon future patient care.

Keywords

Digital teaching files, Dicom image transfer, Privacy, Legislation

Introduction

The Radiological Society of North America (RSNA), Medical Image Resource Center - Teaching File System (TFS), a server-based application was deployed to create Diagnostic Radiology teaching files.

The source of the images for the TFS was Philips IntelliSpace Picture Archive and Communications System (PACS) (version 4.4, Philips Healthcare Informatics Inc., Foster City, USA). The Digital Imaging and Communications in Medicine (DICOM (.dcm)) image was the file type extracted from PACS. DICOM is an internationally recognized image file format that is used by imaging equipment manufacturers and PACS vendors [1]. Using the DICOM format also allowed us to display the teaching file images on a locally hosted, HTML-5, DICOM viewer, outside of the teaching file server.

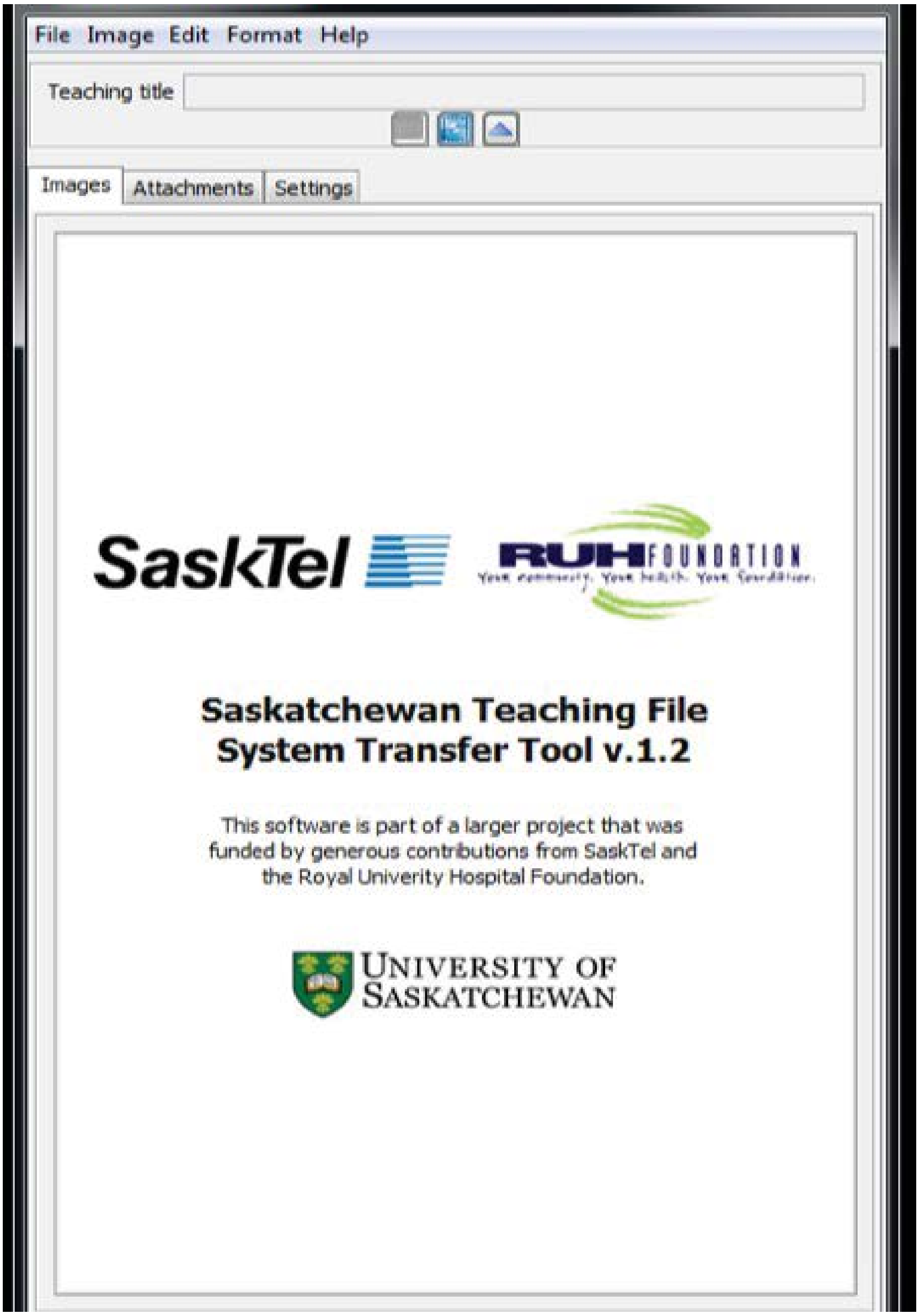
A web-enabled software application was developed and deployed, the Teaching File Transfer Tool (TFTT). The TFTT has DICOM image extraction, anonymization, and editing amalgamated in an integrated software solution (Figure 1). The TFTT communicates with the Philips PACS via an Application Programming Interface (API). The TFTT extracts DICOM images from PACS and first pass anonymizes them at the time of extraction. The user can then review the images and edit them for confidentiality issues, apply annotation(s), crop, and/or blackout portions prior to the revised images being anonymized a second time when they are exported to the TFS.

It is imperative that any strategy used for teaching and research activities safe-guards patient personal health information. This application was approved by our local health authority and the Privacy Commissioner of the province prior to its deployment.

The North American federal legislation pertaining to management of digital patient health information will be discussed in Section 1. Section 2 will discuss the technologies deployed to protect personal health information and develop a successful Diagnostic Radiology teaching file.

Citation: Burbridge B (2020) Dicom Image Anonymization and Transfer to Create a Diagnostic Radiology Teaching File. Int J Radiol Imaging Technol 6:065. doi.org/10.23937/2572-3235.1510065

Accepted: September 16, 2020; **Published:** September 18, 2020


File Image Edit Format Help
Teaching title
Images
Attachments
Settings
SaskTel
RUH FOUNDATION
Your community. Your health. Your foundation.
University of Saskatchewan

Saskatchewan Teaching File System Transfer Tool v.1.2

This software is part of a larger project that was funded by generous contributions from SaskTel and the Royal Univerity Hospital Foundation.

Figure 1: "About TFTT", splash page.

Section 1 - Federal Legislation (North America) Governing the Use of Patient Health Information

Personal information protection and electronic document act (PIPEDA)

This Canadian federal legislation came into effect January 2001. It stipulates how digital personal information should be handled by businesses and by public sector organizations [2]. PIPEDA defines "personal information" as information that describes an identifiable individual. The regulations therefore relate only to information that is directly linked to a person. Health information that has been "de-identified" can be used without obtaining prior consent for secondary purposes such as teaching and research [3].

Cline, et al., in Computer World, analyzed federal health legislation in Canada and the United States [4]. According to Cline, PIPEDA only provides limited guidance on how health information should be de-identified. Principle 5, Section 4.5.3 of Canada's federal privacy law, PIPEDA, states: "Personal information that is no longer required to fulfill the originally identified purposes should be destroyed, erased, or made anonymous.

Table 1: Essential Metadata that must be removed from images anonymized to the HIPAA standard.

1. Names
2. Geographic subdivisions smaller than a state, including address, city, county, precinct, ZIP, and their equivalent geocodes, except for initial three digits of ZIP code
3. All elements of dates (except year) for dates directly related to an individual, including birth date, admission date, discharge date, date of death; and all ages over 89 and all elements of dates (including year) indicative of such age, except that such ages and elements may be aggregated into a single category of age 90 or older
4. Telephone numbers
5. Fax numbers
6. E-mail addresses
7. Social Security numbers
8. Medical record numbers
9. Health plan beneficiary numbers
10. Account numbers
11. Certificate/license numbers
12. Vehicle identifiers and serial numbers, including license plate numbers
13. Device identifiers and serial numbers
14. Web universal locators (URLs)
15. IP address numbers
16. Biometric identifiers, including fingerprints and voice prints
17. Full-face photographic images and any comparable images
18. Other unique identifying numbers, characteristics or codes

Organizations shall develop guidelines and implement procedures to govern the destruction of personal information" [4].

Health insurance portability and accountability act (HIPAA)

HIPAA is federal legislation that covers the use of personal health information in the United States. HIPAA provide useful guidelines on de-identifying health information and it is very specific about how this should be done. HIPAA defines a set of direct identifiers and for a data set to be de-identified it must be stripped of all elements described by the list of direct identifiers (Table 1). These lists include obvious identifiers such as names and medical record numbers, and less obvious identifiers such as service dates and geographic subdivisions smaller than a state. As HIPAA is much more detailed and highly prescriptive, it was the standard for anonymization utilized for this project [5].

Technical Solutions for Safeguarding Patient Information

Image file format

The TFS is capable of storing and displaying images of common file formats: PNG, JPG, DICOM, GIF, etc. To minimize the risk of anonymization problems related to multiple files types, only DICOM images were used.

If images of other formats were felt to be important for a teaching file case, they were converted to the DICOM format by the TFTT, e.g. JPG image converted to DICOM and anonymized.

DICOM header

DICOM images have text-based metadata called a DICOM header which stores tag-value pairs e.g. the tag, "Person Name [0040, A123]" denotes the text value "John Doe". The DICOM tags for an image are standardized internationally, but they can vary somewhat between institutions, modalities, and image sources.

Overview of teaching file creation software

PACSAPIs: The Philips Intellispace 4.4 Picture Archiving and Communication System (PACS) (Foster City, USA) was the primary PACS used to source DICOM images. The central role that the PACS plays in radiology means that teaching file creation must include a mechanism for moving DICOM images from the PACS to the TFS.

A web-based API was developed for the PACS. The primary purpose of the API was to allow PACS users to extract images for handling in the TFTT. The API adds new menu options in PACS that are available when right-clicking on an image (Figure 2). The following PACS Menu operations for the TFTT are displayed by the APIs:

Create a new case. When this option is selected the TFTT is started. A new folder is created for the case on the user's local computer.

Open an existing case. This option allows the user to select a previously created local image folder on their computer. The TFTT is started and the directory path for the case folder is sent to the TFTT.

Export a single DICOM image. The selected DICOM

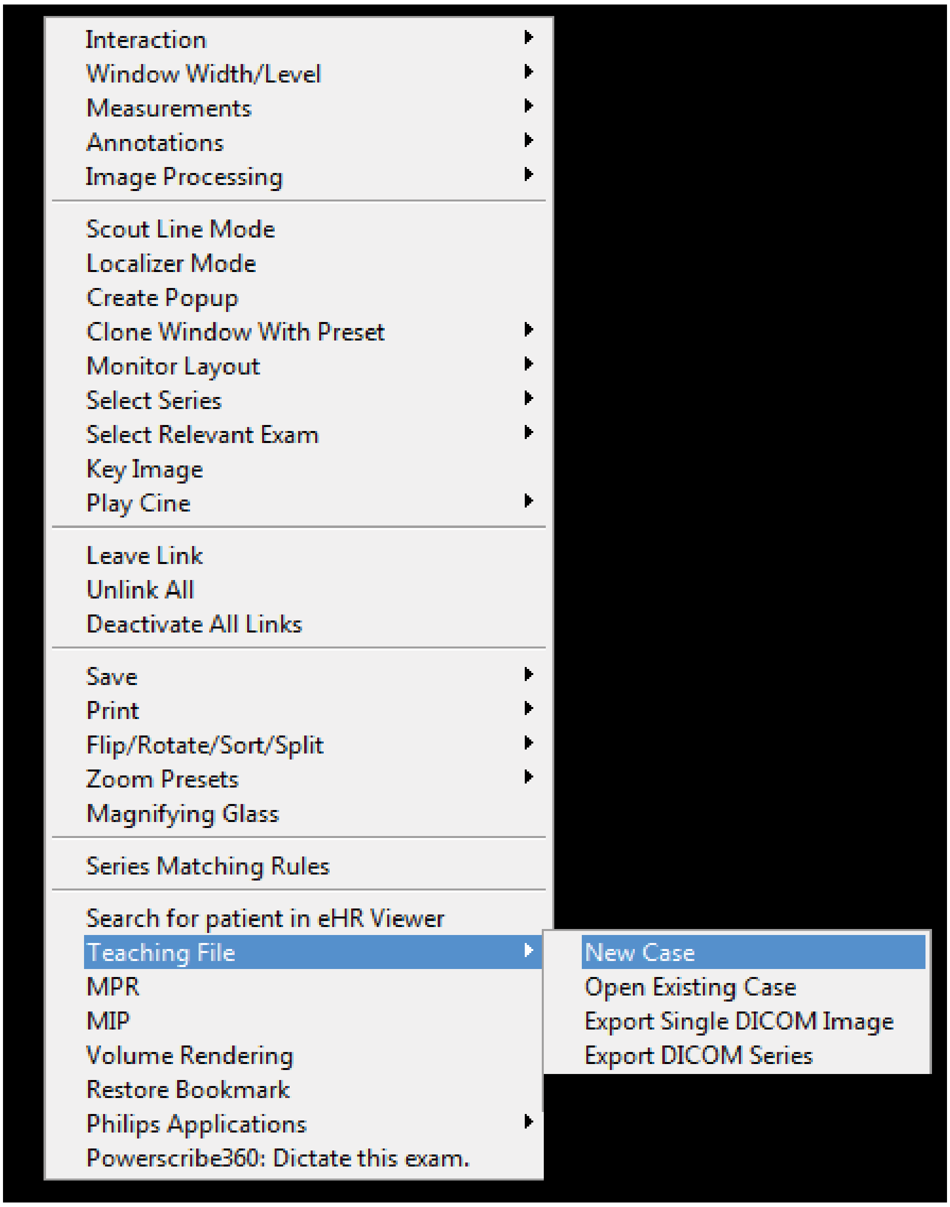


Figure 2: The Teaching pull down menu in PACS.

image is automatically anonymized and written to the current local folder.

Export a DICOM series. The images from a related series i.e. Computed Tomography, Ultrasound, Magnetic Resonance, etc. are written to the current local folder.

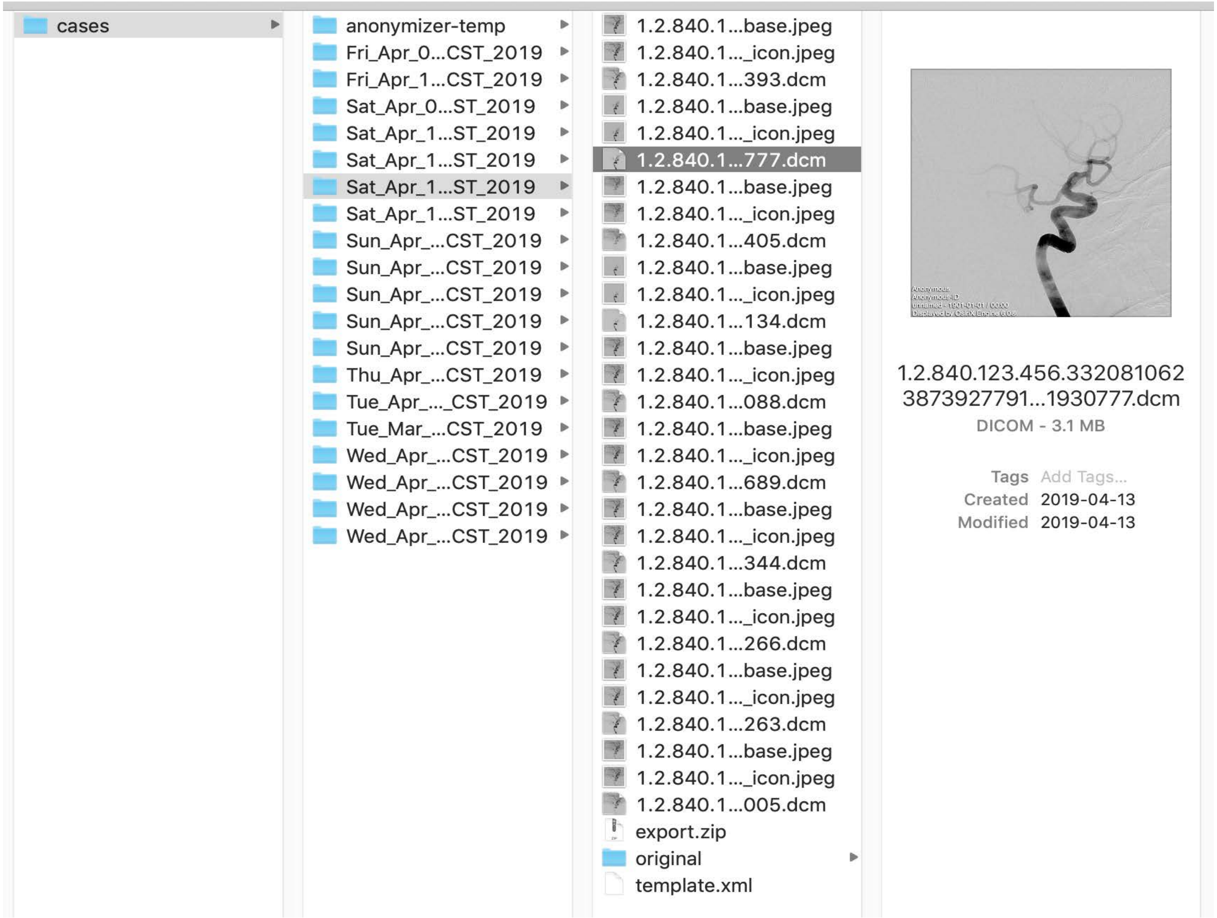
The images transferred for the single image and series image processes are first pass anonymized and held in the local folder. An image of the contents of the local folder is provided in Figure 3. The anonymized images are stored in both DICOM and JPG format in the local folder. After image annotation or editing is completed in the TFTT, the images are second pass anonymized and sent to the TFS.

DICOM image anonymization

The TFTT uses a comprehensive approach for de-identifying DICOM headers: The open-source MIRC - DICOM Anonymizer, developed and released by the RSNA, was integrated into the TFTT. The MIRC Anonymizer is highly customizable, allowing developers to select DICOM tags and to specify how their values should be transformed. For example, tags can be removed, new values can be assigned to tags, or hash functions can be applied to transform their original values.

When images are initially transferred to the TFTT, they are first pass, HIPAA compliant, anonymized, with three exceptions: 1) The station name is preserved, ([0008,1010] Station Name), 2) The institution name is preserved ([0008,0080] Institution Name), and 3) Unique Image Identifiers (UIDs) are preserved. The station name and institution name are used to flag images when de-identification of pixel data is needed. The original UIDs are used by the authoring tool to correctly group and sort sequential images that are part of a series.

Prior to sending the case to the TFS, all images are



- cases

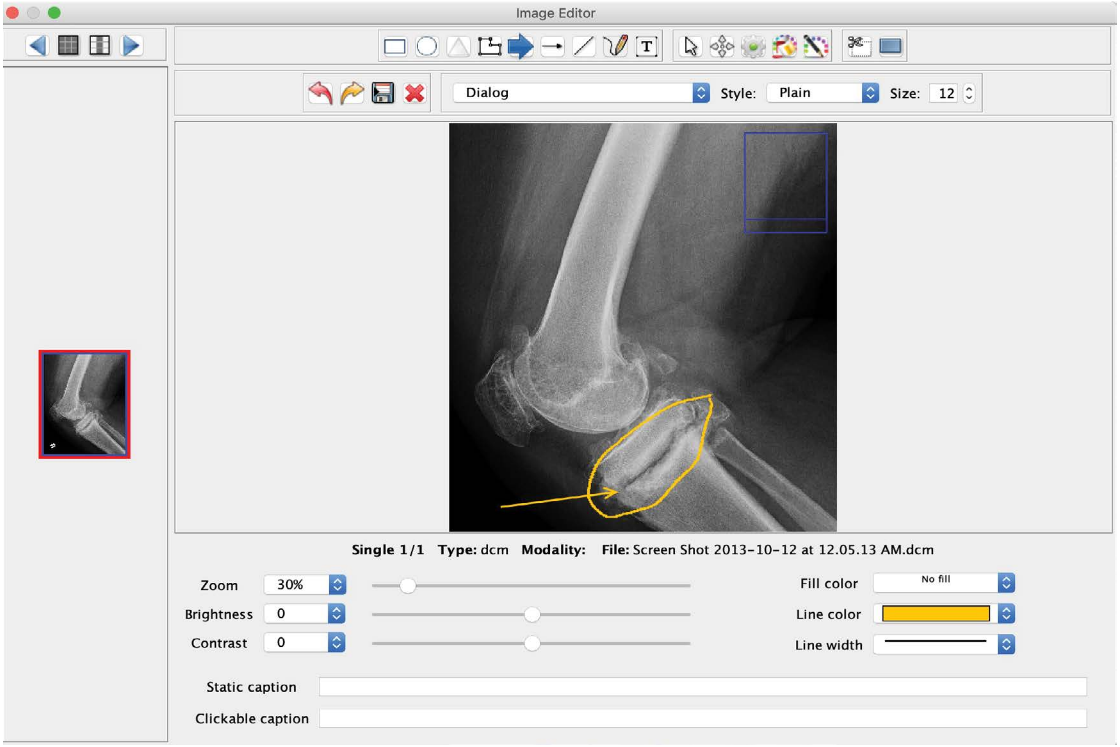
- anonymizer-temp
- Fri_Apr_0...CST_2019
- Fri_Apr_1...CST_2019
- Sat_Apr_0...ST_2019
- Sat_Apr_1...ST_2019
- Sat_Apr_1...ST_2019
- Sat_Apr_1...ST_2019
- Sat_Apr_1...ST_2019
- Sun_Apr_...CST_2019
- Sun_Apr_...CST_2019
- Sun_Apr_...CST_2019
- Sun_Apr_...CST_2019
- Sun_Apr_...CST_2019
- Thu_Apr_...CST_2019
- Tue_Apr_..._CST_2019
- Tue_Mar_...CST_2019
- Wed_Apr_...CST_2019
- Wed_Apr_...CST_2019
- Wed_Apr_...CST_2019
- Wed_Apr_...CST_2019

- 1.2.840.1...base.jpeg
- 1.2.840.1..._icon.jpeg
- 1.2.840.1...393.dcm
- 1.2.840.1...base.jpeg
- 1.2.840.1..._icon.jpeg
- 1.2.840.1...777.dcm
- 1.2.840.1...base.jpeg
- 1.2.840.1..._icon.jpeg
- 1.2.840.1...405.dcm
- 1.2.840.1...base.jpeg
- 1.2.840.1..._icon.jpeg
- 1.2.840.1...134.dcm
- 1.2.840.1...base.jpeg
- 1.2.840.1..._icon.jpeg
- 1.2.840.1...088.dcm
- 1.2.840.1...base.jpeg
- 1.2.840.1..._icon.jpeg
- 1.2.840.1...689.dcm
- 1.2.840.1...base.jpeg
- 1.2.840.1..._icon.jpeg
- 1.2.840.1...344.dcm
- 1.2.840.1...base.jpeg
- 1.2.840.1..._icon.jpeg
- 1.2.840.1...266.dcm
- 1.2.840.1...base.jpeg
- 1.2.840.1..._icon.jpeg
- 1.2.840.1...263.dcm
- 1.2.840.1...base.jpeg
- 1.2.840.1..._icon.jpeg
- 1.2.840.1...005.dcm
- export.zip
- original
- template.xml

1.2.840.123.456.332081062
3873927791...1930777.dcm
DICOM - 3.1 MB

Tags Add Tags...
Created 2019-04-13
Modified 2019-04-13

Figure 3: The local folder holds the images after first and second pass anonymization. An example of the anonymized jpeg and dcm images is provided.



Dialog Style: Plain Size: 12

Single 1/1 Type: dcm Modality: File: Screen Shot 2013-10-12 at 12.05.13 AM.dcm

Zoom 30%
Brightness 0
Contrast 0

Fill color No fill
Line color
Line width

Static caption
Clickable caption

Figure 4: The editing features of the TFTT are illustrated. All of the functions available are displayed in the top of the image i.e. circle, arrow, blackout, crop, etc.

anonymized a second time so that values associated with HIPAA direct identifiers, including those that were preserved during the first pass, are removed. DICOM images often have file names that match the value stored in tag [0008,0018] SOP Instance UID. The anonymizer replaces this value using a hash function, and the file name is changed so that it is identical to the new value.

If anonymization errors occurred with the MIRC Anonymizer, the user was notified, the image was removed from the case and placed in quarantine in the local folder, and a yellow quarantine banner was displayed on the thumbnail icon in the TFTT. The quarantined image cannot be passed to the TFS from the TFTT.

De-identify pixel data: The TFTT includes an image editing feature for removing patient information from image pixel data in DICOM, PNG, JPG, and GIF images.

Embedded pixel data is encountered with any modality that captures and stores images as screen captures. This TFTT feature works with single-frame DICOM images and also allows users to make changes to DICOM multi frame image series by automatically applying image transformations to all frames in the image stack. The TFTT supports the following pixel de-identification operations:

Blackout: Users select a rectangular region in the image, and when the blackout transformation is applied, the color in the selected pixels changes to black.

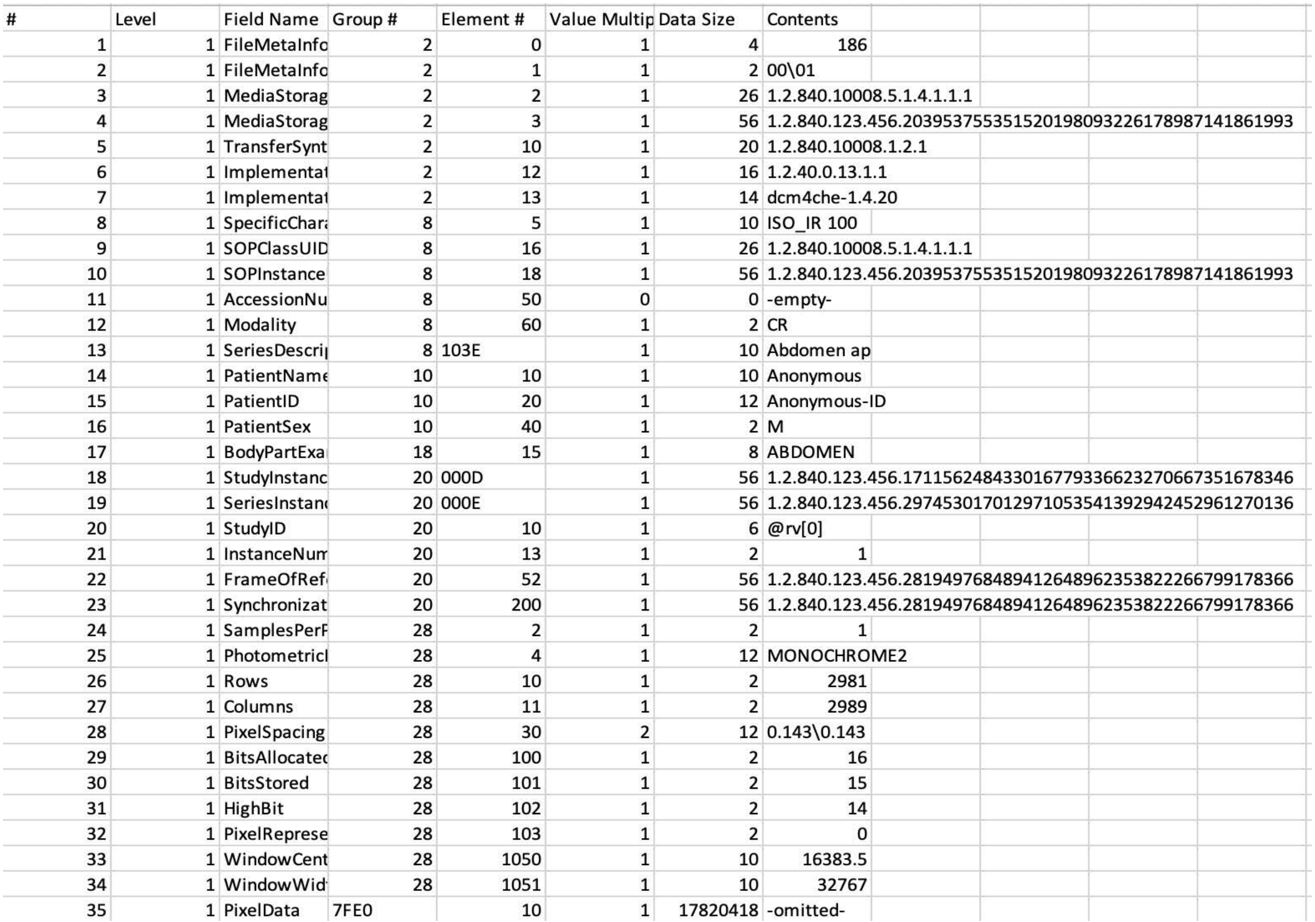
Cropping: Users select part of the image and once the crop transformation is applied the image dimensions change removing all pixels that fall outside of the selected region.

Saving: When transformations have been applied, the user can to save the changes. Once the changes are saved, they become permanent, and the original image cannot be restored. Figure 4 illustrates the editing features of the TFTT (Figure 4).

Validation of image anonymization

Images and image series from the following modalities: radiography, fluoroscopy, angiography, ultrasound, computed tomography and magnetic resonance imaging. Were selected from PACS to assess the effectiveness of the anonymization process.

Firstly, the contents of the original, raw. DICOM header was saved for review. Secondly, the first pass anonymized images were stored in the local folder. Lastly, after image editing in TFTT, the second pass anonymized images were saved for analysis. Hence, the DICOM header information was compared and contrasted for the three unique file types- raw image, first pass anonymization, and second pass anonymization. The DICOM metadata headers were analyzed using an



#	Level	Field Name	Group #	Element #	Value Multip	Data Size	Contents
1	1	FileMetaInfo	2	0	1	4	186
2	1	FileMetaInfo	2	1	1	2	00\01
3	1	MediaStorag	2	2	1	26	1.2.840.10008.5.1.4.1.1.1
4	1	MediaStorag	2	3	1	56	1.2.840.123.456.203953755351520198093226178987141861993
5	1	TransferSynt	2	10	1	20	1.2.840.10008.1.2.1
6	1	Implementat	2	12	1	16	1.2.40.0.13.1.1
7	1	Implementat	2	13	1	14	dcm4che-1.4.20
8	1	SpecificChara	8	5	1	10	ISO_IR 100
9	1	SOPClassUID	8	16	1	26	1.2.840.10008.5.1.4.1.1.1
10	1	SOPInstance	8	18	1	56	1.2.840.123.456.203953755351520198093226178987141861993
11	1	AccessionNu	8	50	0	0	-empty-
12	1	Modality	8	60	1	2	CR
13	1	SeriesDescri	8	103E	1	10	Abdomen ap
14	1	PatientName	10	10	1	10	Anonymous
15	1	PatientID	10	20	1	12	Anonymous-ID
16	1	PatientSex	10	40	1	2	M
17	1	BodyPartExa	18	15	1	8	ABDOMEN
18	1	StudyInstanc	20	000D	1	56	1.2.840.123.456.171156248433016779336623270667351678346
19	1	SeriesInstan	20	000E	1	56	1.2.840.123.456.297453017012971053541392942452961270136
20	1	StudyID	20	10	1	6	@rv[0]
21	1	InstanceNum	20	13	1	2	1
22	1	FrameOfRef	20	52	1	56	1.2.840.123.456.281949768489412648962353822266799178366
23	1	Synchronizat	20	200	1	56	1.2.840.123.456.281949768489412648962353822266799178366
24	1	SamplesPerF	28	2	1	2	1
25	1	PhotometricI	28	4	1	12	MONOCHROME2
26	1	Rows	28	10	1	2	2981
27	1	Columns	28	11	1	2	2989
28	1	PixelSpacing	28	30	2	12	0.143\0.143
29	1	BitsAllocated	28	100	1	2	16
30	1	BitsStored	28	101	1	2	15
31	1	HighBit	28	102	1	2	14
32	1	PixelReprese	28	103	1	2	0
33	1	WindowCent	28	1050	1	10	16383.5
34	1	WindowWid	28	1051	1	10	32767
35	1	PixelData	7FE0	10	1	17820418	-omitted-

Figure 5: An image of the Post-TFTT anonymized DICOM head file for an abdomen X-ray.

open-source DICOM image viewer Horos (GNU Lesser General Public License, Horos, version 3.3 [6].

All HIPAA level DICOM tags were successfully anonymized for both the first pass and second pass image sets. The anonymization process was successful for all modalities investigated (radiography, fluoroscopy, angiography, ultrasound, computed tomography and magnetic resonance imaging). An example of the anonymized DICOM header for an abdomen x-ray has been provided in Figure 5. The raw image DICOM header for this X-ray consisted of 119 tags which were anonymized and diminished to 35 tags.

Discussion

Anonymization of DICOM images is a challenge encountered by those desiring to utilize this file format for teaching and research. Aryanto, et al., highlighted the inadequate performance of a variety of non-commercial DICOM anonymizers that they evaluated [7]. Aryan to evaluated ten non-commercial DICOM anonymizers and found the following, “Only one tool was able to de-identify all required elements with the default setting. Not all of the toolkits provide a customizable de-identification profile. Six tools allowed changes by selecting the provided profiles, giving input through a graphical user interface (GUI) or configuration text file, or providing the appropriate command-line arguments. Using adjusted settings, four of those six toolkits were able to perform full de-identification” [7].

The TFTT allows for successful, HIPAA compliant, two pass anonymization of images from a local version of Philips Intellispace PACS. This functionality is facilitated by the web-enabled application and two desktop APIs that link the TFTT to Philips Intellispace PACS. The anonymized images can then be uploaded to the TFS to create Diagnostic Radiology teaching file cases.

Teaching file cases from our local TFS server, published to the public domain, can be accessed at: https://mistrprodnew.usask.ca:8443/query

A supplemental, HTML-5, DICOM viewer that displays only DICOM images for teaching and learning, can be accessed at: https://mistr.usask.ca/odin/

Both of these teaching resources have robust search tools to interrogate the database for user specific requirements. Individual images, or series of images, can be downloaded for personal use from either site and each teaching case has a unique url that can be embedded in word processing documents (Word, Pages, etc.), PDFs, presentation software (PowerPoint, Keynote, etc.), and web sites, to facilitate teaching and learning. Patient care is enhanced as learners at multiple levels of heath care education (undergraduate and postgraduate) have access to peer-reviewed, anonymized teaching cases with images that can be used for a wide variety of teaching and learning activities.

Declaration of Interest

None.

Funding Sources

The development and deployment of the TFTT, the associated APIs and plugins, the TFS server, the ODIN server, and the software coding for this project was jointly funded by the College of Medicine, University of Saskatchewan and SaskTel via the Royal University Hospital Foundation.

References

1. National Electrical Manufacturers Association (NEMA), The DICOM Standard.
2. The Personal Information Protection and Document Act (PIPEDA). Government of Canada.
3. von Zweck C (2001) Privacy and health information-what are the issues? Occupational Therapy Now 13.1: 25-32.
4. Cline J (2009) Privacy matters: When is personal data truly de-identified? Computer World.
5. Summary of HIPAA privacy rules. U.S. Department of Health and Human Services.
6. Horos DICOM viewer.
7. Aryanto KYE, Oudkerk M, Van Ooijen MPA (2015) Free DICOM de-identification tools in clinical research: Functioning and safety of patient privacy. Eur Radiol 25: 3685-3695.