

# Anonymization of DICOM Files in HTML5 Based Web Browser for Radiation Therapy

P. Rana<sup>1</sup>, W. Sleeman IV<sup>12</sup>, M. Poblacion<sup>1</sup>, J. Palta<sup>12</sup>, P. Ghosh<sup>1</sup> and R. Kapoor<sup>12</sup>

<sup>1</sup> Virginia Commonwealth University, Richmond, US

<sup>2</sup> National Radiation Oncology Program, Veterans Health Affairs, US



## INTRODUCTION

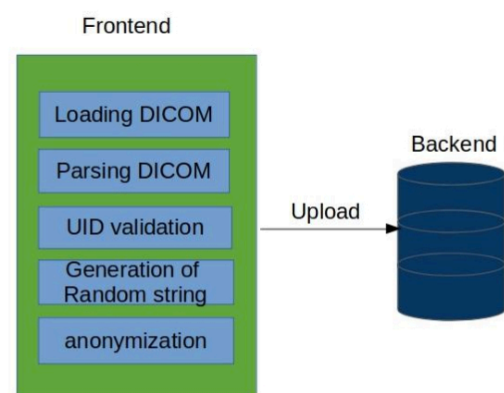
There is a growing need to share DICOM-RT data amongst institutions for rapidly assembling datasets to address practice quality improvements and Big Data science research. Before sharing the data, the files must be anonymized by removing protected health information (PHI) to protect patient privacy. Several desktop installed applications such as DICOMCleaner provide anonymization features but they lack a zero-footprint browser-based functionality which is critical where only facility approved applications can be installed.

## AIM

We propose a prototype zero-footprint HTML5 browser tool that is able to anonymize and upload and multiple DICOM files in a sequential manner where the anonymization task is performed in the browser memory so that no PHI leaves the facility at any instance of time.

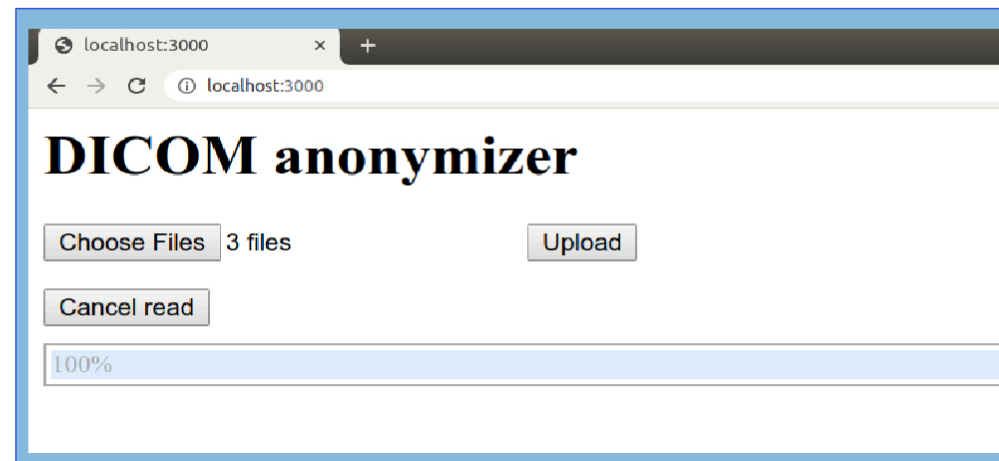
## METHOD

- DICOM anonymization tool is built on the popular dicomParser JavaScript library.
- We have parsed out the sensitive DICOM tags and overwritten them with the new anonymized value of the same length or emptied the tag values completely.
- The pipeline of this tool consists of DICOM upload, DICOM parsing, UID validation, anonymization and sending anonymized DICOM to the backend.



## RESULTS

Our prototype system has the ability to anonymize 40 different tags covering 19 standard PHI item groups. Some of these removed/modified tags are Study Date, Series Date, Study Time, Series Time, Institution Name, Institution Address, Referring Physicians Name, Referring Physicians Address, Referring Physicians Telephone Number, Operators Name, Patients Name, Patient ID, Study ID, Patients Address, Date Time, Person Name, Patient's Sex, Patient's Birth Date, Study Instance UID and others. The tool also creates fresh unique identifiers (UIDs) and consistently replaces them in the datasets preserving the established relationships between the files. Our tool also supports multiple file uploads so that a complete dataset of DICOM files can be anonymized and uploaded in a single session. It takes about 20 ms per file and total time grows linearly with the number of files due to increased computation.



GUI of Dicom Anonymization tool

File Number	Time (ms)
1	19.63
5	98.52
10	200.1
20	401.3
50	1003.2
100	2005.2

Time for Dicom Anonymization

Attribute	Original	De-identified	send
Patient Name	Patient_00001	h7elosWYBELu	
Patient ID	00001	8d1MN	
Patient Birth Date		20200627	
Study Description	PROSTATE	Z2npkEH8	
Protocol Name	Onco Pelvis Large2/OncoBody	iibJ27DRXJhwnsZYqSobpiHFJuB	
Accession #			
Study Id	22881	2bwYs	
Study Date	20100616	20200627	
Study Time		192351	
Series Description			
Acquisition Date	20100616	20200627	
Acquisition Time		192351	
Content Date	20100616	20200627	
Content Time		192351	

Original & anonymized value

## CONCLUSIONS

For the first time, we designed a prototype that anonymizes multiple DICOM files and uploads it to a secure server. Our tools currently anonymize DICOM tags with the same length of new values, as rewriting a longer value requires resizing the byte array and updating the length of the elements, which can be added later. Our tool shows great promises to anonymize DICOM tags besides preserving the relationship between DICOM UIDs. This tool has some limitations; for example, it does not detect the PHI details in the free text of original data. However, in the future, an NLP approach can be used to identify PHI in the free-text. Another possible improvement can be giving the user the option to select which tags they want to anonymize.

## REFERENCES

<https://github.com/cornerstonejs/dicomParser>

Monteiro, E., Costa, C., & Oliveira, J. L. (2017). A de-identification pipeline for ultrasound medical images in DICOM format. *Journal of medical systems*, 41(5), 89.

Newhauser, W., Jones, T., Swerdloff, S., Newhauser, W., Cilia, M., Carver, R., ... & Zhang, R. (2014). Anonymization of DICOM electronic medical records for radiation therapy. *Computers in biology and medicine*, 53, 134-140.

González, D. R., Carpenter, T., van Hemert, J. I., & Wardlaw, J. (2010). An open source toolkit for medical imaging de-identification. *European radiology*, 20(8), 1896-1904.

Suzuki, H., Amano, M., Kubo, M., Kawata, Y., Niki, N., & Nishitani, H. (2007, March). Anonymization server system for DICOM images. In *Medical Imaging 2007: PACS and Imaging Informatics* (Vol. 6516, p. 65160Z). International Society for Optics and Photonics.

Riddle, W. R., & Pickens, D. R. (2005). Extracting data from a DICOM file. *Medical physics*, 32(6Part1), 1537-1541.

Aryanto, K. Y. E., Oudkerk, M., & van Ooijen, P. M. A. (2015). Free DICOM de-identification tools in clinical research: functioning and safety of patient privacy. *European radiology*, 25(12), 3685-3695.

## ACKNOWLEDGEMENTS

The work was funded by the VHA's National Radiation Oncology Program office under contract with Virginia Commonwealth University as part of the Health Information Gateway and Exchange (HINGE) development project.

## CONTACT INFORMATION

William Sleeman

william.sleemaniv@vcuhealth.org