UNIVERSITY OF OTTAWA
HEART INSTITUTE
INSTITUT DE CARDIOLOGIE
DE L'UNIVERSITÉ D'OTTAWA

Practical Approaches to De-identification of Medical Imaging Data

From Metadata to Pixel-level Protection

Nadia Farrag, Ph.D.

November 25, 2025

Please note:

Any use of software highlighted in this presentation must comply with individual institutional policies for software use.

UNIVERSITY OF OTTAWA
HEART INSTITUTE
INSTITUT DE CARDIOLOGIE
DE L'UNIVERSITÉ D'OTTAWA

Agenda & Learning Objectives

Topic	Learning Objectives
Introduction	Understand why de-identification is essential
Legal frameworks	Learn key requirements under HIPAA & PHIPEDA
What's in an image?	Distinguish metadata vs pixel content
DICOM basics	Recognize important tags + Supplement 142 rules
Metadata workflows	Learn how to remove/transform PHI in metadata
Pixel workflows	Understand burned-in text removal + defacing
Tools & demos	Compare viewers, GUI tools, and coding options
Quality Control	Intro to essential QC checks + assess re-ID risk
Summary	Apply final tips & identify next steps/resources

Introduction

UNIVERSITY OF OTTAWA HEART INSTITUTE
INSTITUT DE CARDIOLOGIE DE L'UNIVERSITÉ D'OTTAWA

Introduction

- More imaging research → More image sharing
- Open science = greater need for proper de-identification
- Goal → safely prepare imaging data for reuse & sharing

UNIVERSITY OF OTTAWA
HEART INSTITUTE
INSTITUT DE CARDIOLOGIE
DE L'UNIVERSITÉ D'OTTAWA

Purpose of De-identification

- Protect patients
- Comply with federal/provincial laws
- Enable safe data sharing (e.g., TCPS2 waiver of consent)

Legal Frameworks

UNIVERSITY OF OTTAWA
HEART INSTITUTE
INSTITUT DE CARDIOLOGIE
DE L'UNIVERSITÉ D'OTTAWA

Legal frameworks

PIPEDA

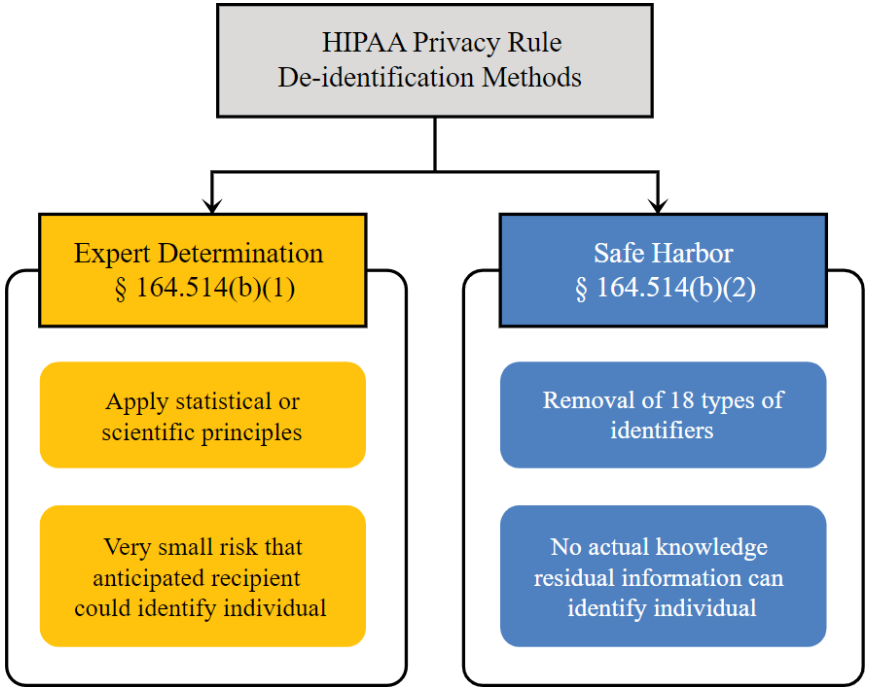
HIPAA

GDPR

UNIVERSITY OF OTTAWA HEART INSTITUTE
INSTITUT DE CARDIOLOGIE DE L'UNIVERSITÉ D'OTTAWA

HIPAA Compliance

- Two options for de-identification
  - Expert determination
  - Safe harbor
- Safe Harbor Method is the most commonly used
  - List of 18 specific identifiers that must be removed to consider data de-identified



UNIVERSITY OF OTTAWA HEART INSTITUTE
INSTITUT DE CARDIOLOGIE DE L'UNIVERSITÉ D'OTTAWA

HIPAA Compliance

1. Name
2. Address
3. Dates
4. Telephone number
5. Fax number
6. Email address
7. Social security number
8. Medical record number
9. Health beneficiary number
10. Account number
11. Certificate/ license number
12. Vehicle identifiers
13. Device identifiers/ serial numbers
14. Web URLs
15. IP address
16. Biometric identifiers
17. Full face photos
18. Any other unique identifier

UNIVERSITY OF OTTAWA
HEART INSTITUTE
INSTITUT DE CARDIOLOGIE
DE L'UNIVERSITÉ D'OTTAWA

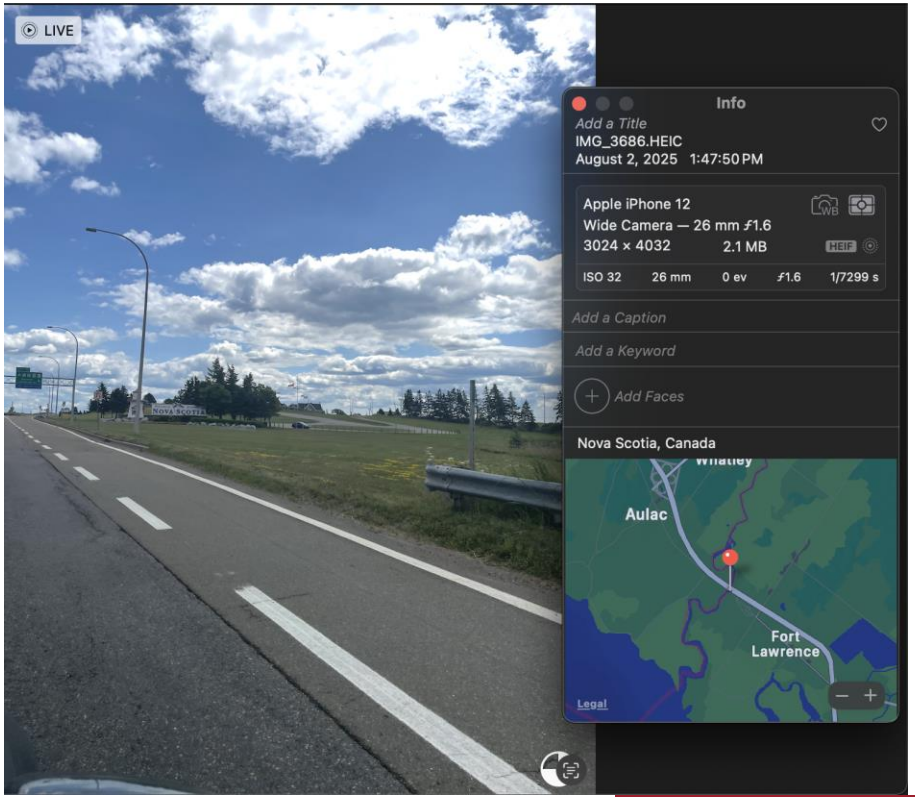
PHIPEDA Compliance

- No fixed list of identifiers
- Remove/transform direct + indirect identifiers
- Ensure very low re-ID risk

What's in an image?

UNIVERSITY OF OTTAWA
HEART INSTITUTE
INSTITUT DE CARDIOLOGIE
DE L'UNIVERSITÉ D'OTTAWA

What's in an image?

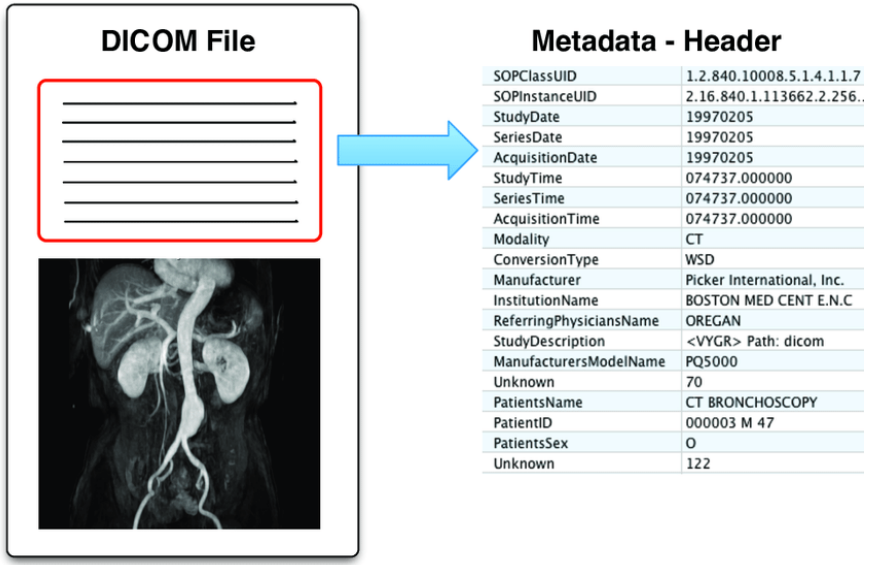


- Pixel data (visible component)
- Meta-data (invisible text component)
- Text component not directly visible but accessible
- Applicable to all image file types (e.g., JPEG, PNG, TIFF, DICOM)

DICOM Format

UNIVERSITY OF OTTAWA HEART INSTITUTE
INSTITUT DE CARDIOLOGIE DE L'UNIVERSITÉ D'OTTAWA

DICOM Format



- Standard format for medical imaging
- Combines pixel data and hundreds of header tags in a single file
  - Meta-data (text) component cannot be separated from the image
- Tags contain dates, IDs, UIDs, equipment info
- First published in 1992 (ISO 12052)

UNIVERSITY OF OTTAWA HEART INSTITUTE
INSTITUT DE CARDIOLOGIE DE L'UNIVERSITÉ D'OTTAWA

DICOM Format

- Each DICOM tag is composed of:
  - *Tag number*: in the format (XXXX,XXXX)
  - *Value Representation* (VR): describes the data type
  - *Value length:* defines the length of the attribute's value in bytes
  - *Value field:* the actual data

Tag	VR	Data	Length	Description
0054,0016	SQ		1	Radiopharmaceutical Information Sequence
>0018,1071	DS	0	1	Radiopharmaceutical Volume
>0054,0300	SQ		1	Radionuclide Code Sequence
>>0008,0100	SH	Tc-99m	6	Code Value
>>0008,0102	SH	99SDM	5	Coding Scheme Designator
0054,0400	SH	32	2	Image ID

UNIVERSITY OF OTTAWA HEART INSTITUTE
INSTITUT DE CARDIOLOGIE DE L'UNIVERSITÉ D'OTTAWA

DICOM Format

- There are over 2,000 defined tags
- New tags added as modalities evolve
- All official tags are listed in **DICOM Part 6: Data Dictionary:** https://dicom.nema.org/medical/dicom/current/output/pdf/part06.pdf
- A searchable list of DICOM tags can be found here: https://www.dicomlibrary.com/dicom/dicom-tags/

UNIVERSITY OF OTTAWA HEART INSTITUTE
INSTITUT DE CARDIOLOGIE DE L'UNIVERSITÉ D'OTTAWA

DICOM Format: Standard & Private Tags

- **Standard** = universal & documented
  - E.g.: Dates/Times, UIDs, etc.
- **Private** = vendor-specific & inconsistent
- May be defined in manufacturer manuals, but not always
- Vendors often store acquisition details inside private tags
  - E.g.: b-values for diffusion weighted imaging MRI

1. A sample list of Private Attributes thought to be safe is provided here. Vendors do not guarantee them to be safe, and do not commit to sending them in any particular software version (including future products).

Data Element	Private Creator	VR	VM	Meaning
(7053,xx00)	Philips PET Private Group	DS	1	SUV Factor – Multiplying stored pixel values by Rescale Slope then this factor results in SUVbw in g/l
(7053,xx09)	Philips PET Private Group	DS	1	Activity Concentration Factor – Multiplying stored pixel values by Rescale Slope then this factor results in MBq/ml.
(00E1,xx21)	ELSCINT1	DS	1	DLP
(01E1,xx26)	ELSCINT1	CS	1	Phantom Type
(01E1,xx50)	ELSCINT1	DS	1	Acquisition Duration
(01F1,xx01)	ELSCINT1	CS	1	Acquisition Type
(01F1,xx07)	ELSCINT1	DS	1	Table Velocity
(01F1,xx26)	ELSCINT1	DS	1	Pitch
(01F1,xx27)	ELSCINT1	DS	1	Rotation Time
(0019,xx23)	GEMS_ACQU_01	DS	1	Table Speed [mm/rotation]

UNIVERSITY OF OTTAWA HEART INSTITUTE
INSTITUT DE CARDIOLOGIE DE L'UNIVERSITÉ D'OTTAWA

DICOM Format: Structured Reports

- Text-based clinical report encoded in DICOM
- May link to images or measurements
- Must be checked for PHI

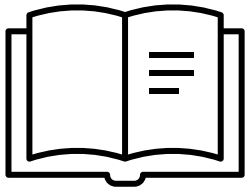
De-identification of DICOM data: Meta-data workflows

UNIVERSITY OF OTTAWA HEART INSTITUTE
INSTITUT DE CARDIOLOGIE DE L'UNIVERSITÉ D'OTTAWA

Meta-data De-Identification in DICOM

- Remove or replace obvious tags containing PHI (e.g., Patient Name, Patient ID, Study Date, etc.)
- Some tags contain PHI but are not as obvious – researcher's discretion

+



DICOM supplement 142

+

Researcher's discretion

UNIVERSITY OF OTTAWA HEART INSTITUTE
INSTITUT DE CARDIOLOGIE DE L'UNIVERSITÉ D'OTTAWA

DICOM Meta-data: Tags to look out for

- Dates galore – there is more than one tag for date!
- Most UIDs other than SOP Class UID
  - SOP Instance UID (0002, 0003)
  - Media Storage SOP Instance UID (0008, 0013)
  - Series Instance UID (0020, 000E)
  - etc.
- Private tags
- Free-text fields

UNIVERSITY OF OTTAWA HEART INSTITUTE
INSTITUT DE CARDIOLOGIE DE L'UNIVERSITÉ D'OTTAWA

DICOM Meta-data: Tags to look out for

Field Name	Tag	Content
StudyDate	(0008,0020)	20251104 ←
SeriesDate	(0008,0021)	20251104 ←
StudyTime	(0008,0030)	120323.000000
SeriesTime	(0008,0031)	120632.453999
MediaStorageSOPInstanceUID	(0002,0003)	1.3.12.2.1107.5.2.18.41999.30000025110412545190100000104
SeriesInstanceUID	(0020,000E)	1.3.12.2.1107.5.2.18.41999.30000025110412545190100000102
Unknown	(0021,1060)	20251104120629

UNIVERSITY OF OTTAWA HEART INSTITUTE
INSTITUT DE CARDIOLOGIE DE L'UNIVERSITÉ D'OTTAWA

DICOM Meta-data: Tags to look out for

- DICOM Supplement 142 is an appendix to the DICOM standard
- Letter-code system – guidelines to delete, empty, zero out, replace, etc.

X: Remove attribute

K: Keep as-is

U: Generate new UID

… etc.

Full list here: https://dicom.nema.org/medical/dicom/current/output/html/part15.html#table_E.1-1
Full DICOM supplement 142 can be found here: https://www.dicomstandard.org/News-dir/ftsup/docs/sups/sup142.pdf

UNIVERSITY OF OTTAWA HEART INSTITUTE
INSTITUT DE CARDIOLOGIE DE L'UNIVERSITÉ D'OTTAWA

Summary of DICOM Supplement 142

Category	What Must Be Removed / Replaced	How
Direct Identifiers	Patient name, IDs, birth date/time, contact info, family info, comments	Remove or replace (Z/D)
People & Places	All staff names/IDs; institution name/address; station/department/AE titles	Remove or replace (X/D)
Dates, Times, UIDs	All dates/times; all Study/Series/SOP/Frame UIDs and referenced UIDs	Remove/shift dates; replace UIDs (U)
Private & Free Text	All private tags; any text fields that can hold identifiers (descriptions, comments, protocol names)	Remove or clean (X/C)
Graphics & Pixels	Overlays/curves/annotations; burned-in text; recognizable facial features	Remove or clean; blackout pixels if needed

UNIVERSITY OF OTTAWA
HEART INSTITUTE
INSTITUT DE CARDIOLOGIE
DE L'UNIVERSITÉ D'OTTAWA

How to De-identify DICOM Data: Three Practical Pathways

1. Using viewing software

2. Using dedicated de-identification tools

3. Programming / Coding libraries

TECHNICAL COMPLEXITY

TECHNICAL COMPLEXITY

TECHNICAL COMPLEXITY

DICOM De-identification Using Viewing Software

UNIVERSITY OF OTTAWA HEART INSTITUTE
INSTITUT DE CARDIOLOGIE DE L'UNIVERSITÉ D'OTTAWA

DICOM De-identification Using Viewing Software

Mac

HOROS

OSIRIX

Windows

RadiAnt DICOM VIEWER

Windows

Mac

Linux

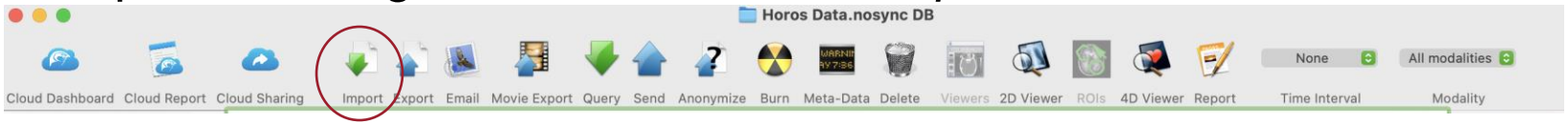
WEASIS medical viewer

Mac

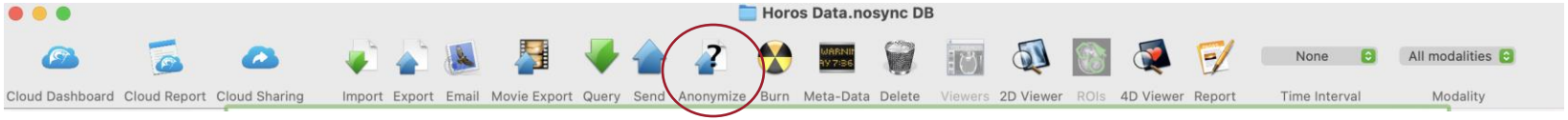
LOW
TECHNICAL COMPLEXITY

DICOM De-identification Using HOROS

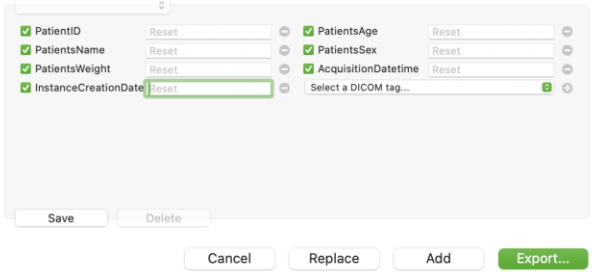
1. Import the image file or folder under *Import*



2. Hit *Anonymize* and select the relevant DICOM tags in the pop-up window



Pop up window:



- "Select a DICOM tag" has a drop-down function & allows user to customize which tags to remove (+)
- *Export* saves the de-identified study in a new folder

Mac

TECHNICAL COMPLEXITY

DICOM De-identification Using HOROS

UNIVERSITY OF OTTAWA
HEART INSTITUTE
INSTITUT DE CARDIOLOGIE
DE L'UNIVERSITÉ D'OTTAWA

DICOM De-identification Using Viewing Software

Mac

HOROS

OSIRIX

Windows

RadiAnt DICOM VIEWER

Windows

Mac

Linux

WEASIS medical viewer

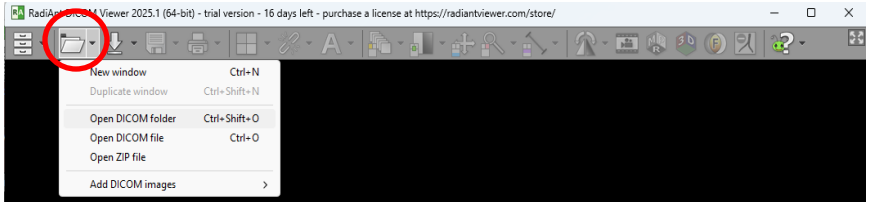
Windows

LOW

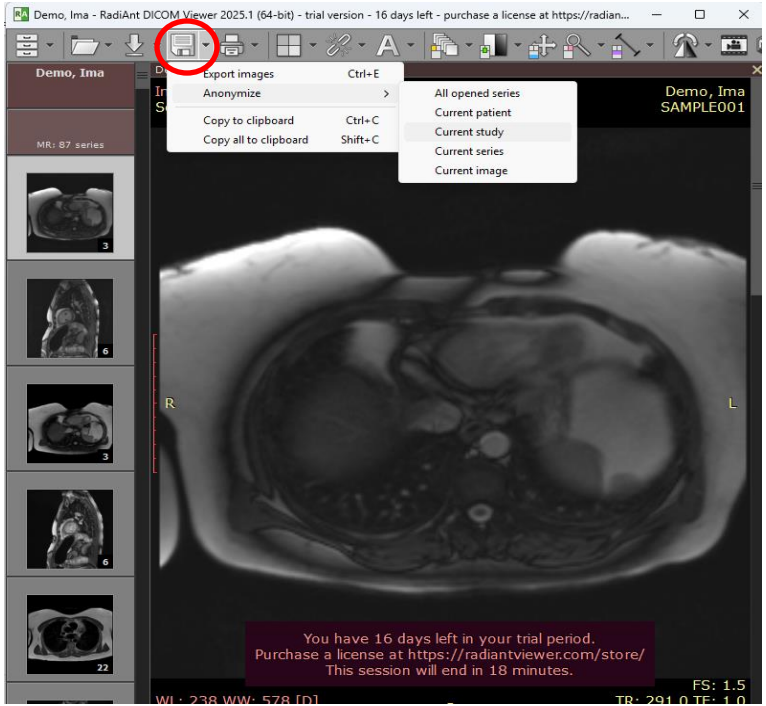
TECHNICAL COMPLEXITY

DICOM De-identification Using RadiAnt

1. Import the image file or folder under *Import*



2. Under *Save* drop-down arrow, navigate to *Anonymize*, and select relevant option

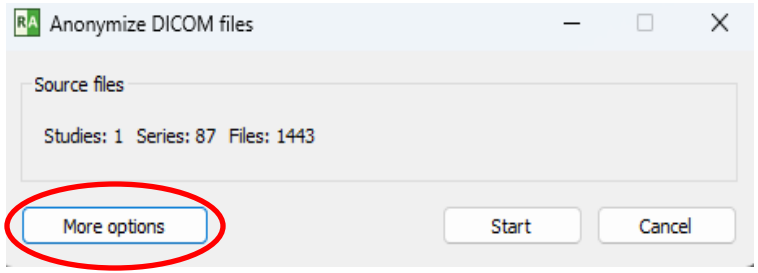


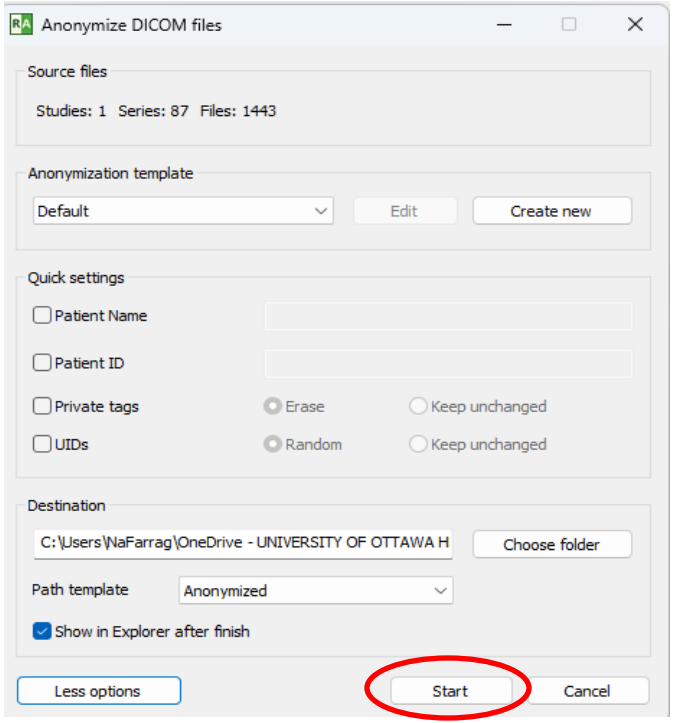
Windows

LOW
TECHNICAL COMPLEXITY

DICOM De-identification Using RadiAnt

3. In the pop-up, choose *Start* to de-identify the image under the default options OR hit *More options* to customize





UNIVERSITY OF OTTAWA HEART INSTITUTE
INSTITUT DE CARDIOLOGIE DE L'UNIVERSITÉ D'OTTAWA

DICOM De-identification Using Viewing Software

Pros

- User control
- GUI based – no coding required
- Good for small datasets
- Good for demos / education
- Easy to catch odd tags

Cons

- Human error
- Slow, not scalable
- No error logs
- DICOM structure learning cuve
- 
  Limited ability for pixel removal

Dedicated DICOM De-identification Tools – download online, use in a graphical user interface (GUI)

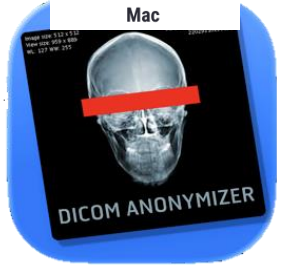
UNIVERSITY OF OTTAWA HEART INSTITUTE
INSTITUT DE CARDIOLOGIE DE L'UNIVERSITÉ D'OTTAWA

Dedicated DICOM De-Identification Tools

Windows Mac

CARE DICOM Anonymizer

Mac



DICOM Anonymizer

Windows Mac Java

RSNA®

Clinical Trial Processor (CTP)

Windows Mac Java

PixelMed Publishing™

DICOMCleaner

Windows Mac Java

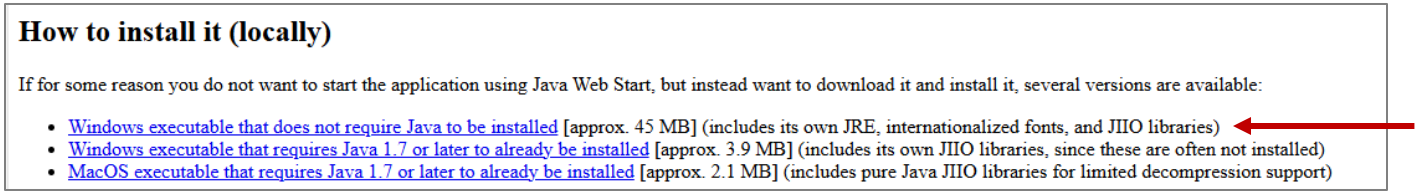
NITRC Neuroimaging Tools & Resources Collaboratory

TECHNICAL COMPLEXITY

DICOM Deidentification on PixelMed DICOMCleaner

1. Click this link for full installation directions: How to use DicomCleaner™

2. Scroll to *How to install it (locally)* & choose the appropriate option



How to install it (locally)

If for some reason you do not want to start the application using Java Web Start, but instead want to download it and install it, several versions are available:

- Windows executable that does not require Java to be installed [approx. 45 MB] (includes its own JRE, internationalized fonts, and JIIO libraries)
- Windows executable that requires Java 1.7 or later to already be installed [approx. 3.9 MB] (includes its own JIIO libraries, since these are often not installed)
- MacOS executable that requires Java 1.7 or later to already be installed [approx. 2.1 MB] (includes pure Java JIIO libraries for limited decompression support)

Direct links if needed:

https://www.dclunie.com/pixelmed/software/webstart/DicomCleaner.html (Windows)

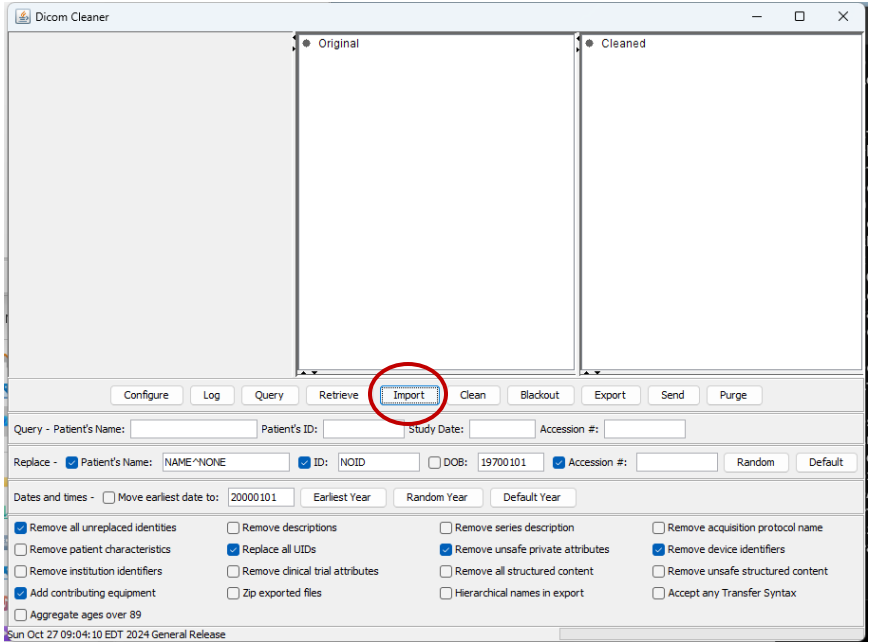
https://www.osirix-ukusergroup.org/dicom-cleaner (Mac*)

*You will also need Java to run this program on Mac.

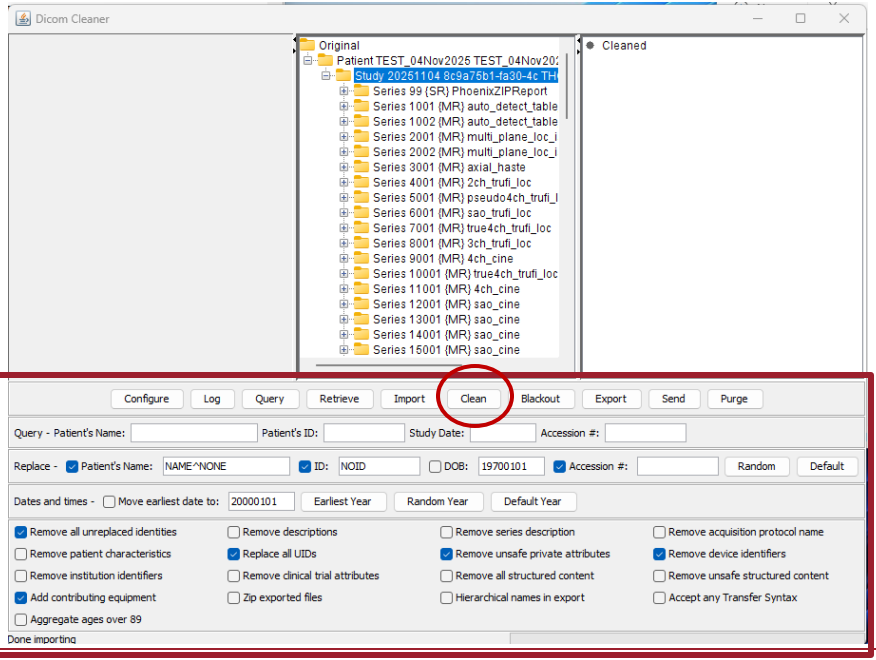
TECHNICAL COMPLEXITY

DICOM Deidentification on PixelMed DICOMCleaner

3. Hit *Import* and navigate to image directory



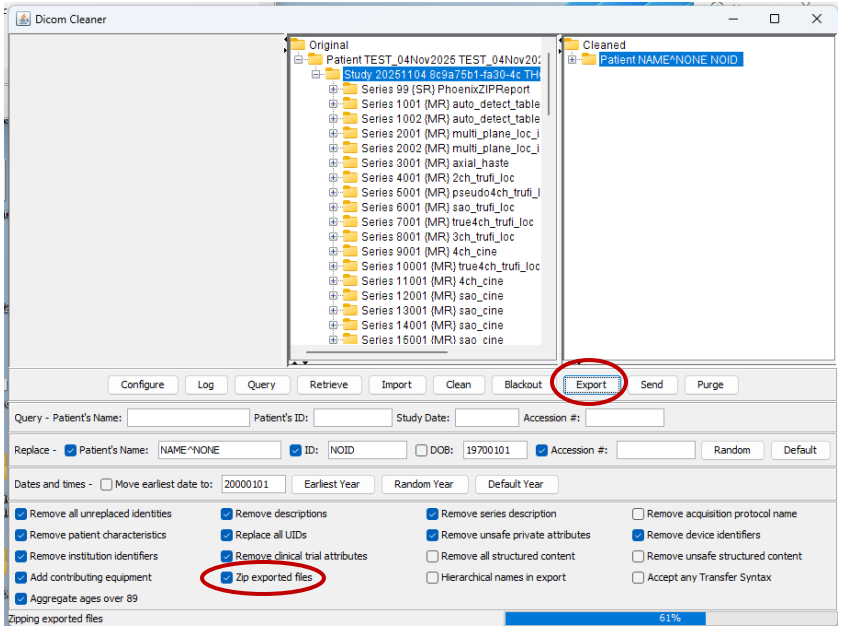
4. Select tags to deidentify in the control panel and hit *Clean*



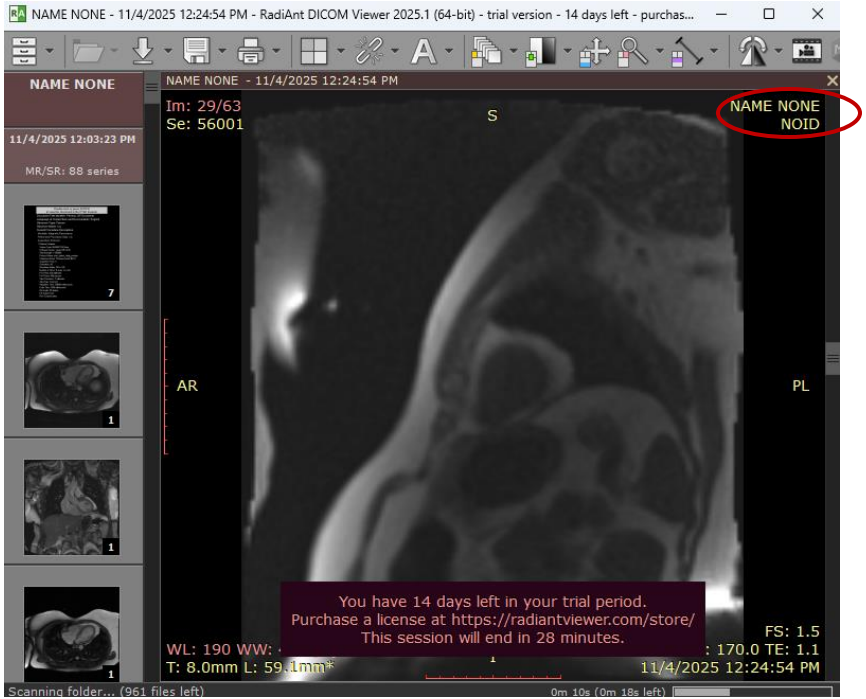
TECHNICAL COMPLEXITY

DICOM Deidentification on PixelMed DICOMCleaner

4. Decide if you want to compress the exported files (checkbox in control panel) and hit *Export*



5. Verify the output in viewing software

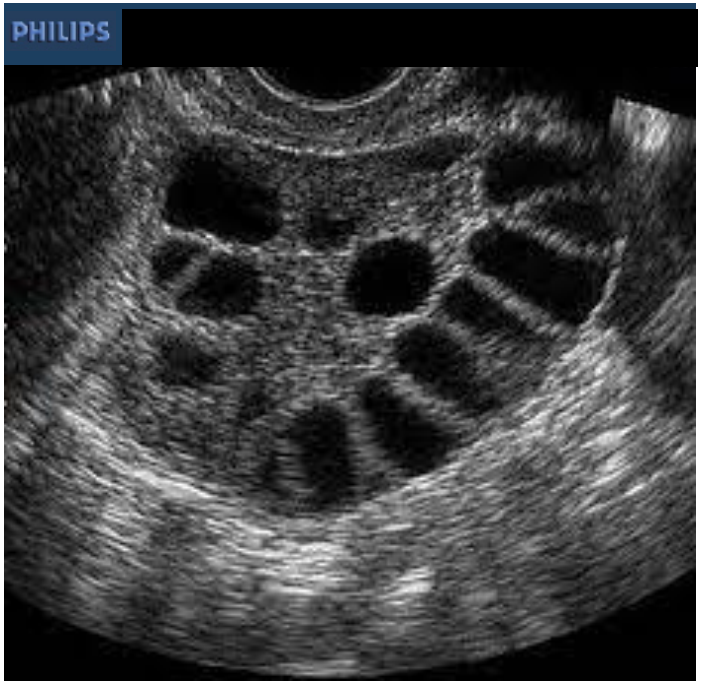


De-identification of DICOM data: Pixel-level workflows

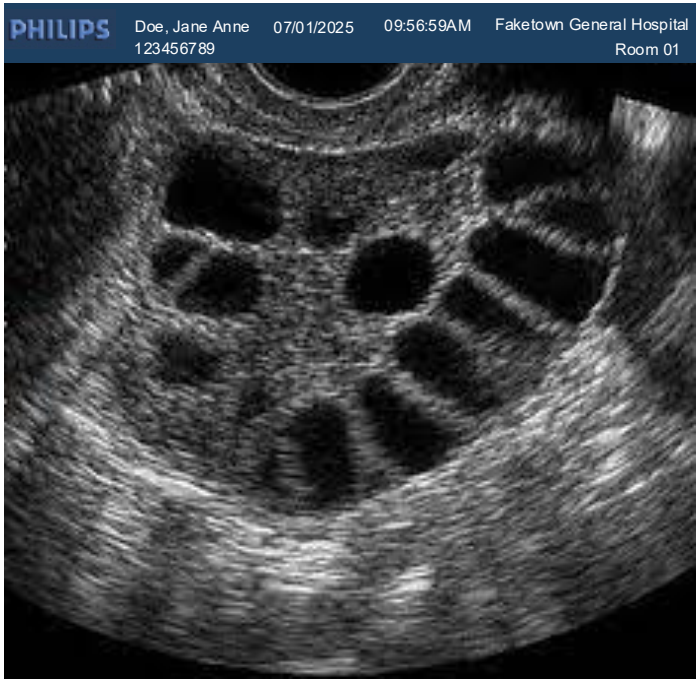
UNIVERSITY OF OTTAWA HEART INSTITUTE
INSTITUT DE CARDIOLOGIE DE L'UNIVERSITÉ D'OTTAWA

Pixel-level de-identification of DICOM

Black out PHI



Crop out PHI

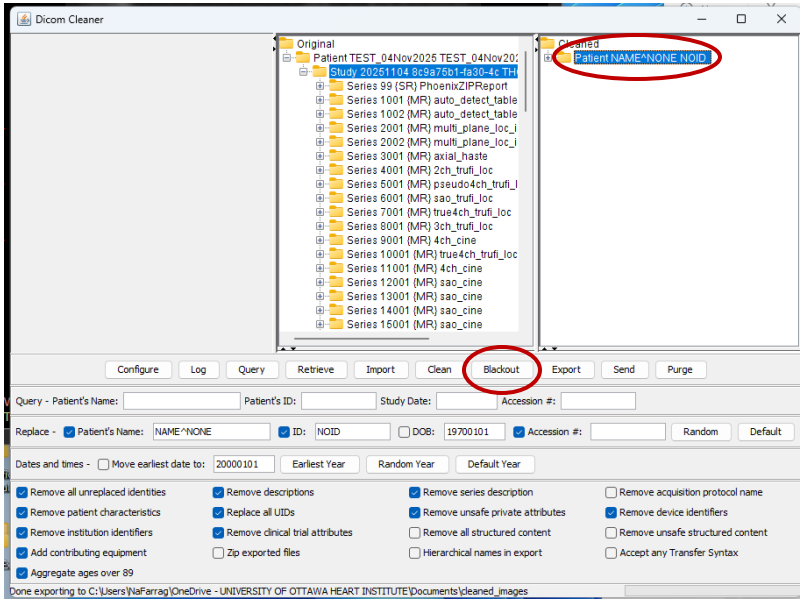


MEDIUM

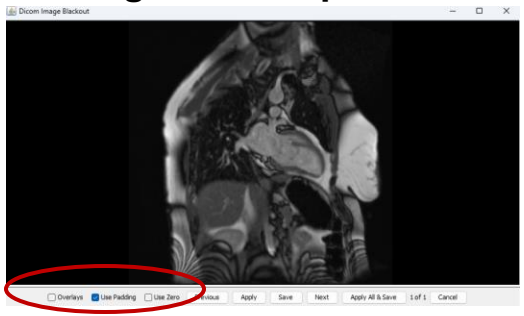
TECHNICAL COMPLEXITY

Black out pixels on PixelMed DICOMCleaner

1. After meta-data deidentification (Step 4 previous slide), under “Cleaned” panel, select your study and hit *Blackout*



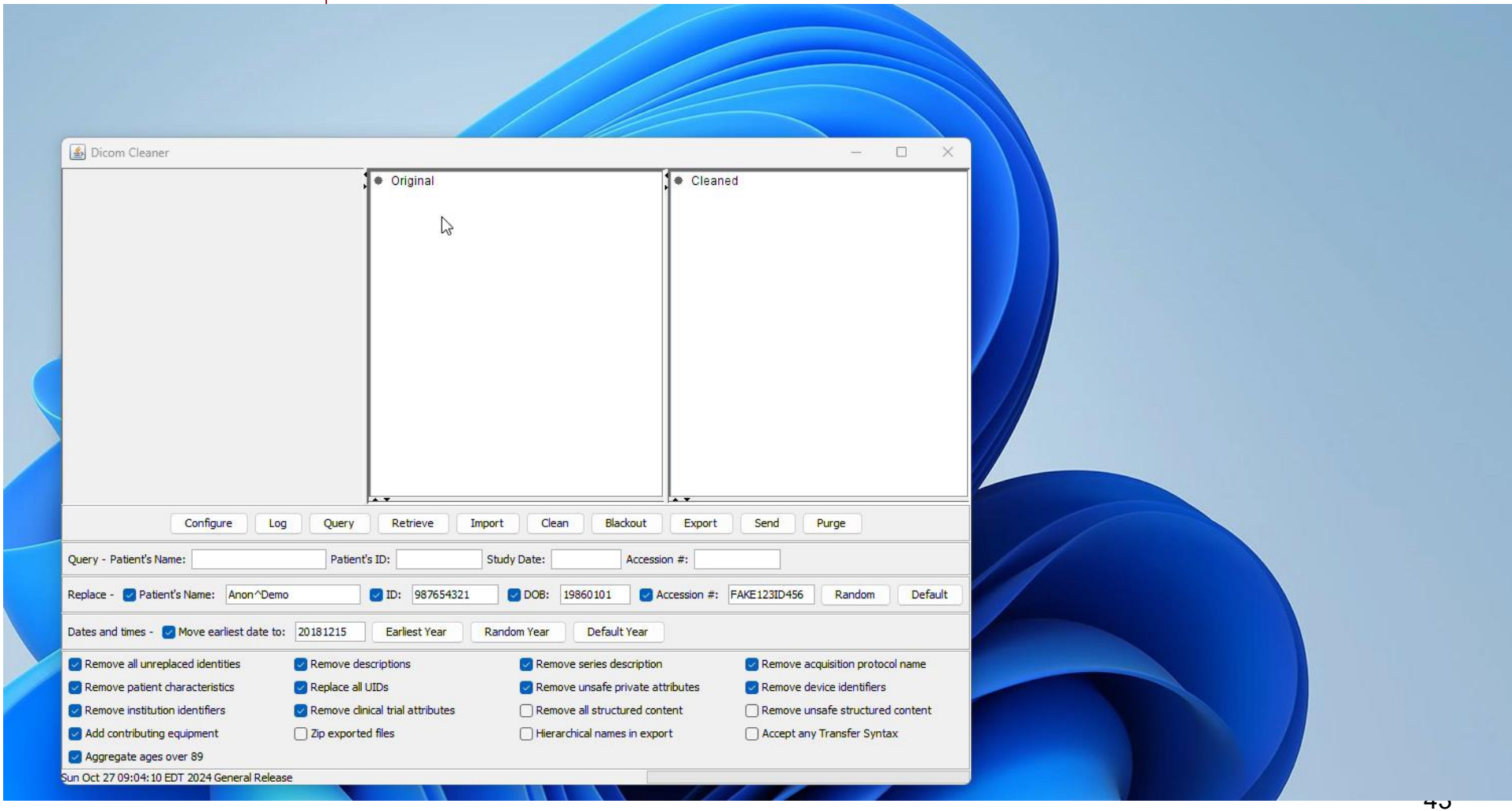
2. In the pop-up window, choose options (padding / zeros)
Use **Right mouse pointer** to select the area for blackout.



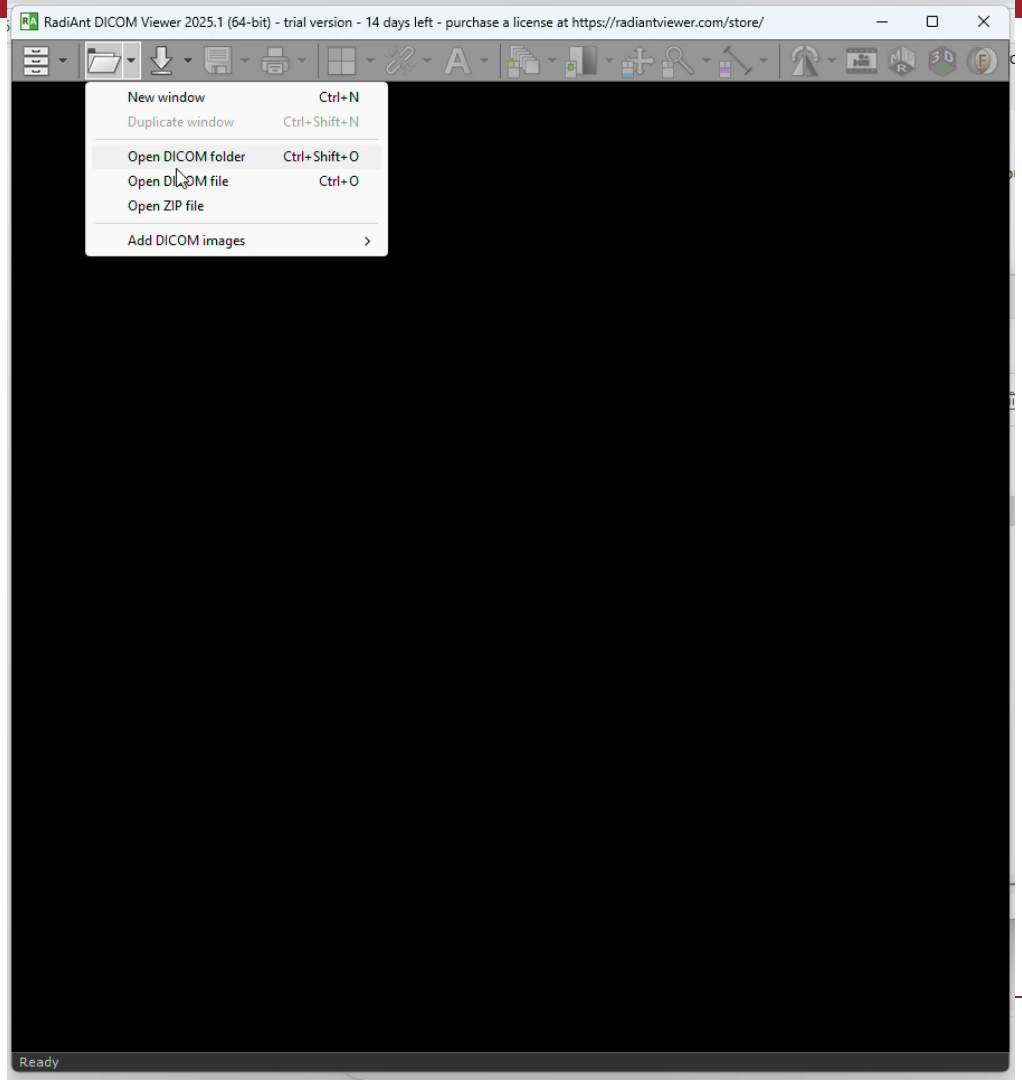
3. After selecting the region for blackout, hit *Apply*

4. Be sure to hit *Save.* If working with multi-slice images requiring the same blackout box, hit *Apply all & Save*

The blacked-out versions stay associated with the entries in the Cleaned panel – pixel PHI is now covered and images are ready to be exported

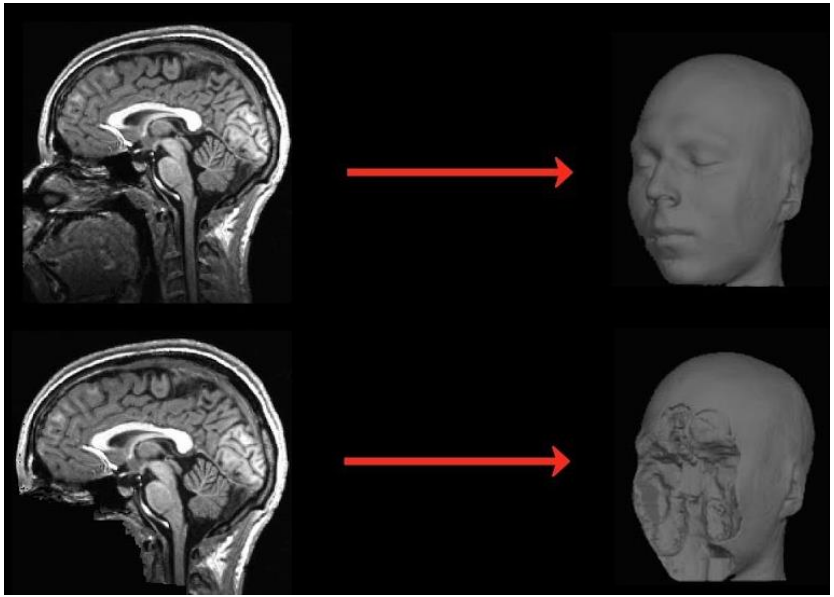

Dicom Cleaner
Original
Cleaned
Configure
Log
Query
Retrieve
Import
Clean
Blackout
Export
Send
Purge
Query - Patient's Name:
Patient's ID:
Study Date:
Accession #:
Replace -
Patient's Name:
Anon^Demo
ID:
987654321
DOB:
19860101
Accession #:
FAKE123ID456
Random
Default
Dates and times -
Move earliest date to:
20181215
Earliest Year
Random Year
Default Year
Remove all unreplaced identities
Remove descriptions
Remove series description
Remove acquisition protocol name
Remove patient characteristics
Replace all UIDs
Remove unsafe private attributes
Remove device identifiers
Remove institution identifiers
Remove clinical trial attributes
Remove all structured content
Remove unsafe structured content
Add contributing equipment
Zip exported files
Hierarchical names in export
Accept any Transfer Syntax
Aggregate ages over 89
Sun Oct 27 09:04:10 EDT 2024 General Release

UNIVERSITY OF OTTAWA
HEART INSTITUTE
INSTITUT DE CARDIOLOGIE
DE L'UNIVERSITÉ D'OTTAWA



TECHNICAL COMPLEXITY

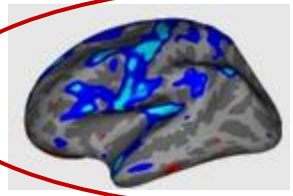
Pixel-level de-identification of DICOM: Defacing



- Removes identifiable facial anatomy
- Can be achieved using algorithmic or AI-based tools
- Some GUIs available to mask pixels, but real defacing demands code-based tools that operate on the 3D anatomy → you will need to use some form of coding/programming

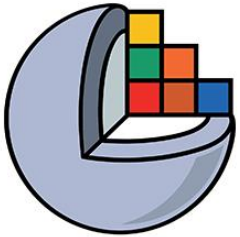
TECHNICAL COMPLEXITY

Pixel-level de-identification of DICOM: Defacing



freesurfer

MRIcroGL



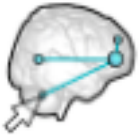
3D Slicer

+ BRAINSdeface extension*

*not stable on newer version of slicer

BRAIN INNOVATION

Brain Voyager



BioImage Suite Web

fast & portable image analysis

TECHNICAL COMPLEXITY

Defacing using Freesurfer

freesurfer set up and use requires coding

1. Download freesurfer here: https://surfer.nmr.mgh.harvard.edu/fswiki/DownloadAndInstall

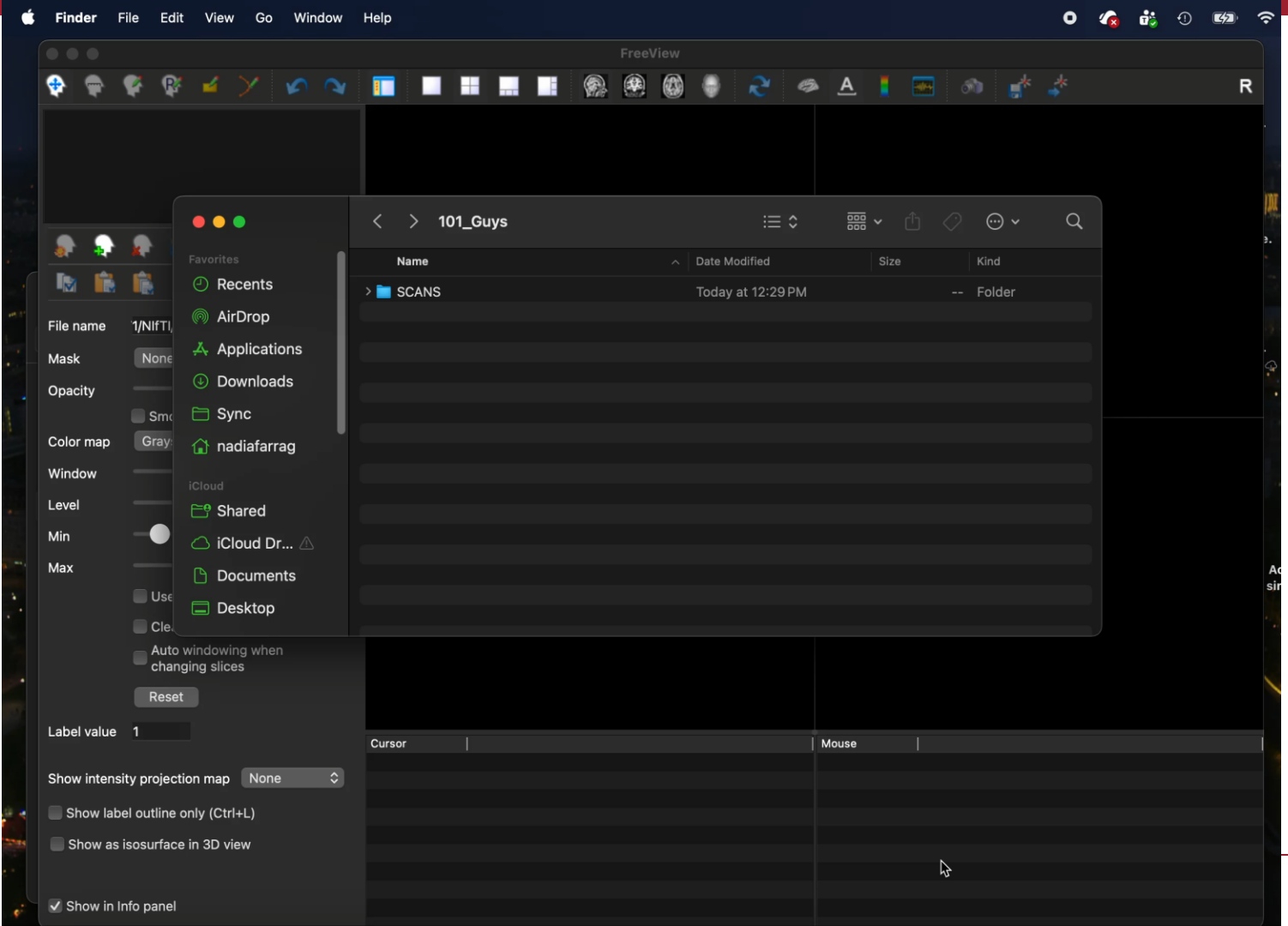
2. Follow instructions to obtain a free license key here: https://surfer.nmr.mgh.harvard.edu/registration.html

3. Initialize freesurfer in command line terminal:

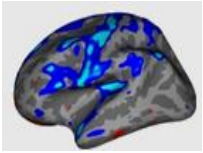
```
export FREESURFER_HOME=/Applications/freesurfer/8.1.0
source $FREESURFER_HOME/SetUpFreeSurfer.sh
```

4. Run the defacing command:

```
mri_deface \
  /path/to/input_image.nii.gz \
  $FREESURFER_HOME/average/talairach_mixed_with_skull.gca \
  $FREESURFER_HOME/average/face.gca \
  /path/to/output_defaced.nii.gz
```



TECHNICAL COMPLEXITY



freesurfer

Mac

Linux

UNIVERSITY OF OTTAWA HEART INSTITUTE
INSTITUT DE CARDIOLOGIE DE L'UNIVERSITÉ D'OTTAWA

Deidentification of DICOM data using Dedicated Tools

Pros

- Built-in tools
- Supports some batch processing
- Reduced operator error
- GUI-based – no coding
- Follows DICOM Sup142 standards

Cons

- Functions vary by software
- May require prior configuration
- Manual – user must "hit go"
- Outdated interfaces
- Limited with private tags

De-identification of DICOM data using programming / coding libraries

TECHNICAL COMPLEXITY

Deidentification of DICOM data using programming / coding libraries

PYDICOM

DICOGNITO

deid

mri_deface
pydeface
pyfacewipe

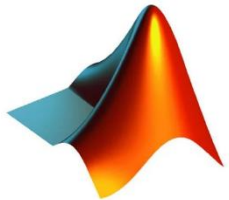
python

DCMTK

GDCM

dicom-anonymizer

Command line interface (CLI)



MATLAB

dicomanon

TECHNICAL COMPLEXITY

Deidentification of DICOM data using programming / coding libraries

Pros

- Fast and scalable
- Automated and reproducible
- Easy batch processing
- Consistent results
- Less operator error

Cons

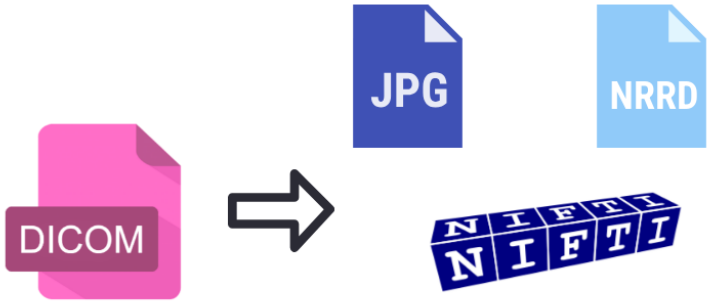
- Small mistakes scale fast
- Learning curve
- Multiple libraries/tools required
- Environment conflicts
- Susceptible to bugs and errors

General Recommendations & Quality Control

UNIVERSITY OF OTTAWA HEART INSTITUTE
INSTITUT DE CARDIOLOGIE DE L'UNIVERSITÉ D'OTTAWA

General Recommendations for De-identification of Medical Images

- Coding workflows scale best for large datasets.
  - If new to coding, online support is available; look into institutional support
- Sharing large datasets: convert from DICOM to alternate
- Conversion is helpful, but not a replacement for de-identification



UNIVERSITY OF OTTAWA HEART INSTITUTE
INSTITUT DE CARDIOLOGIE DE L'UNIVERSITÉ D'OTTAWA

Quality Control of De-identification

Verify PHI removed from meta-data **and** pixel data

Export meta-data to tabular format for easier checks

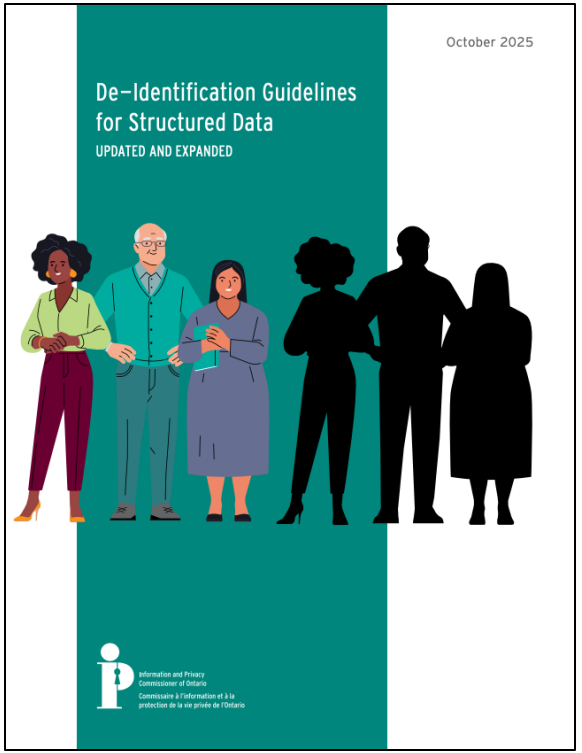
Use OCR to check for burned-in text

Manual review still required

UNIVERSITY OF OTTAWA HEART INSTITUTE
INSTITUT DE CARDIOLOGIE DE L'UNIVERSITÉ D'OTTAWA

Risk of Re-identification

- Meta-data risk measurable with k-anonymity–type models
- Pixel risk → recognition tests (human or AI)
- OCR and/or ML failure rate can be used to measure residual risk



Summary

UNIVERSITY OF OTTAWA
HEART INSTITUTE
INSTITUT DE CARDIOLOGIE
DE L'UNIVERSITÉ D'OTTAWA

Summary and Final Tips

- De-identification in medical imaging = metadata + pixel considerations
- GUI for demos and/or small datasets; coding scales
- Choose the simplest tool that meets your dataset's needs
- Always verify outputs visually + via metadata inspection
- Protect privacy while keeping data scientifically useful

UNIVERSITY OF OTTAWA HEART INSTITUTE
INSTITUT DE CARDIOLOGIE DE L'UNIVERSITÉ D'OTTAWA

References & Resources

- DICOM Standard: https://dicom.nema.org
- HIPAA Privacy Rule: https://www.hhs.gov/hipaa/for-professionals/special-topics/de-identification/index.html
- PIPEDA: https://www.priv.gc.ca/en/privacy-topics/privacy-laws-in-canada/the-personal-information-protection-and-electronic-documents-act-pipeda/
- HOROS Viewer: https://horosproject.org
- RadiAnt Viewer: https://www.radiantviewer.com/
- RSNA CLP: https://wiki.rsna.org/display/CONDET/CTP
- PixelMed DICOMCleaner: https://www.dclunie.com/pixelmed/software/webstart/DicomCleanerUsage.html
- Pydicom: https://pydicom.github.io
- Dicognito: https://github.com/pydicom/dicognito
- Python deid: https://github.com/pydicom/deid

Thank you for your attention

Questions are welcome