Data anonymization at the image level, preserving the original DICOM format of the file

Software for Automatic Anonymization of Radiographic Images

Rugilė DAULIŪTĖ[1], **Eimantas** MAČIUS[1], **Gediminas** DANYS[2], **JUSTAS** Trinkūnas[2,3], **Roma** PURONAITĖ[2,4,5], **Rolandas** BĖRONTAS[2]

INTRODUCTION

- With the entry into force of The General Data Protection Regulation (GDPR), there is a strong focus on the anonymization of medical images and the protection of patient data.
- The main goal is to quickly and qualitatively depersonalize images.
- One of the biggest challenges – sensitive information embedded on the picture pixels.

METHODS

- Programming language used - "Python 3.7"
- Graphical user interface modelling was done with the "Tkinter" framework
- For text recognition "OpenCV" models were used
- Image manipulation was done with "Pillow" framework
- DICOM files manipulation was done using "PyDICOM" library
- Final build was compiled into executable Windows file

RESULTS

- Our software automatically detects text using machine learning algorithms and covers it.
- Our software outputs the result files in a convenient data structure for easy user experience.
- The anonymization is done sufficiently to correspond to current DICOM standards.
- 1 picture takes about 5s to anonymize.

DISCUSSION

- After anonymization – the actual image does not lose quality, so it could be further analyzed.
- We plan to improve the anonymization algorithm and adapt the software to other types of radiological images, including multi-frame images.

Modality	CR	US	CT	DX	MR	XA	ES	RF	NM	SR	OT	XC	IO	PT	SC	OP	PR	MG	PX
%	28.9%	27.8%	12.4%	10.8%	5.4%	4.6%	4.0%	2.6%	1.0%	0.8%	0.7%	0.3%	0.2%	0.1%	0.1%	<0.1%	<0.1%	<0.1%	<0.1%

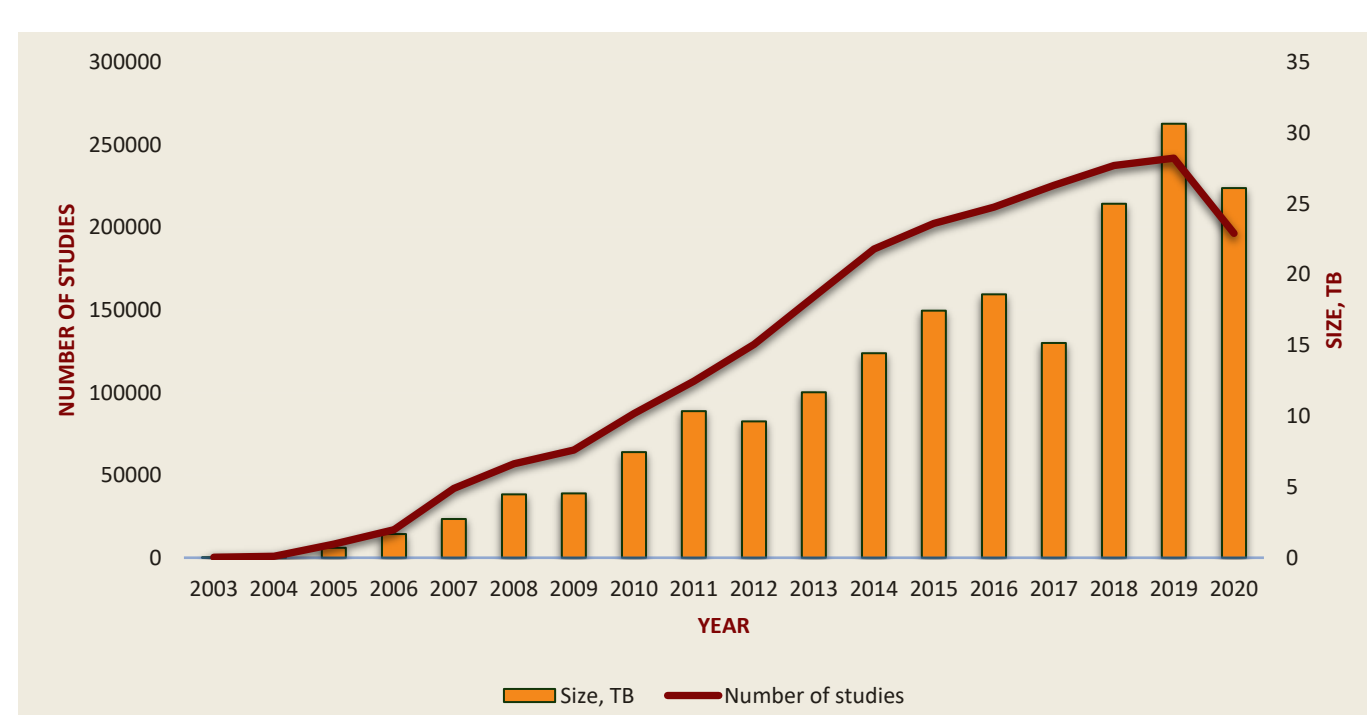**Table 1.** Distribution of studies by imaging modality in Vilnius University Hospital Santaros Klinikos (VUHSK).

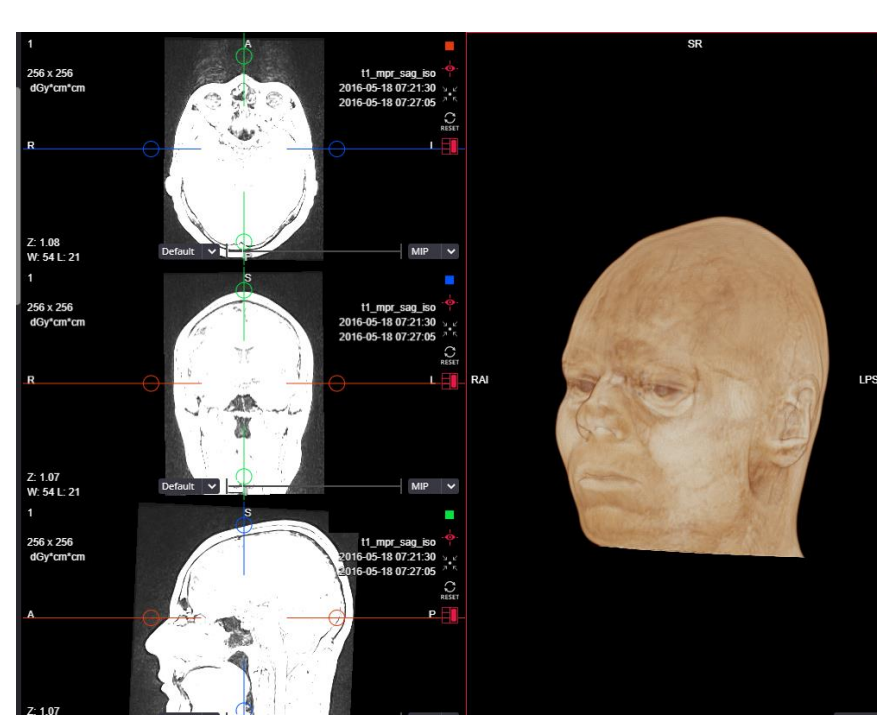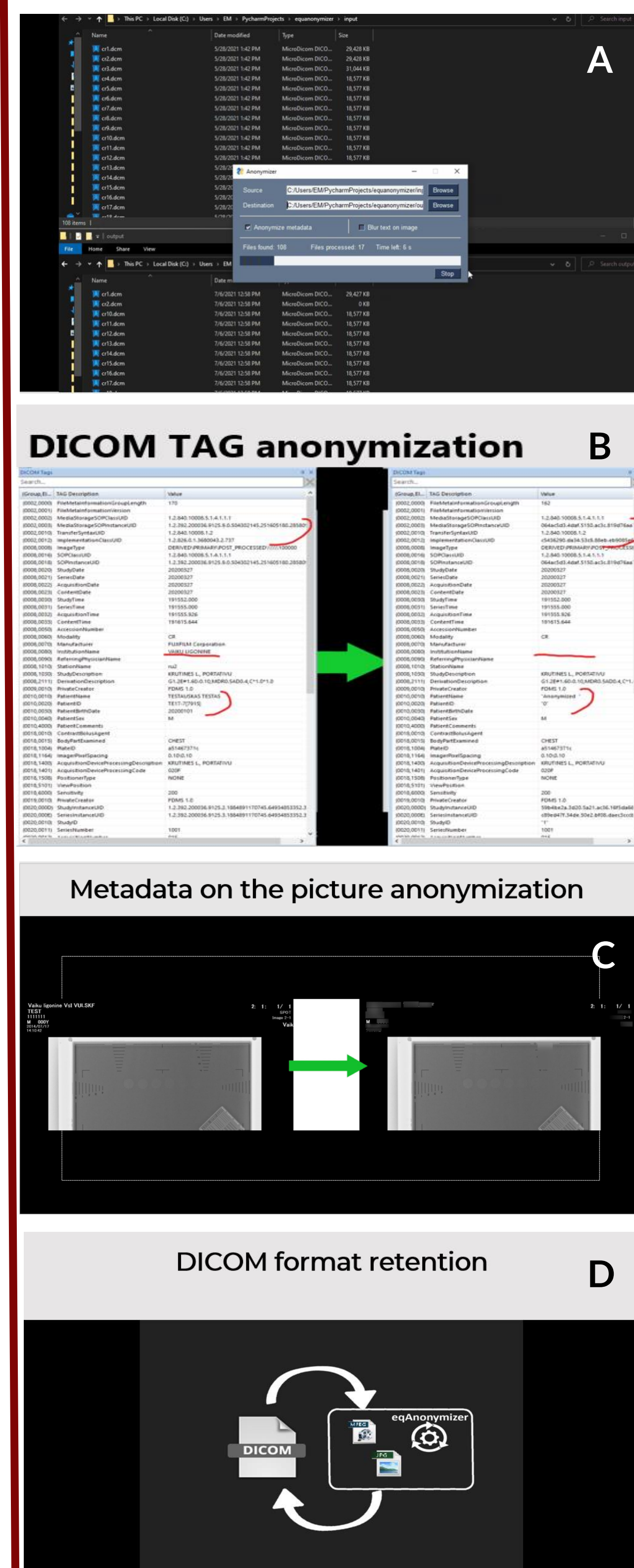**Figure 1.** Total number and size of studies per year in VUHSK.

Figure 2. MRI 3D reconstruction.

How can I anonymize my MR image?

Figure 3. Process of automatic anonymization: (A) initiation, (B) DICOM TAG anonymization, (C) picture anonymization, (D) DICOM format retention.

[1] Innovation incubator "KTU Startup Space" of Kaunas Technology University (MB "Innovatio veterinariae", Registration code 305361185)
[2] Vilnius University Hospital Santaros Klinikos
[3] Vilnius Gediminas Technical University
[4] Institute of Data Science and Digital Technologies, Vilnius University
[5] Faculty of Medicine, Vilnius University
E-mail: roma.puronaite@santa.lt

Take a picture for further information