scientific data

Check for updates

OPEN

ARTICLE

From theory to practice: Re-identification Challenge to test imaging data pseudonymization effectiveness

Rocío Catalán Flores[1]✉, Ignacio Gómez-Rico Junquero[1], Paula Jiménez Gómez[1], Joao Gomes Carvalho[2], Simone Mazzetti[3], Ricard Martínez Martínez[4], José Damián Segrelles Quilis[5], Pau Lozano Lloret[5], Adrián Galiana-Bordera[1], Irene Bargellini[3], Daniele Regge[3], Manuela França[2] & Luis Martí-Bonmatí[1,6]

DICOM image pseudonymization is an effective measure to protect patient privacy and ensure compliance with the GDPR. However, no standardized methods guarantee the irreversible anonymization of DICOM images or provide evidence on the robustness of these procedures. Organizations such as NEMA propose pseudonymization profiles for DICOM metadata, but the risk to data protection is assumed by the entity responsible for pseudonymization, as the potential privacy risks associated with their use cannot be accurately assessed. In the Re-identification Challenge, the selected 68 participants were tasked with re-identifying 38 pseudonymized DICOM studies from multiple European acquisition sites, modalities and anatomical regions. The data were pseudonymized using the pseudonymization profiles developed in ChAImeleon[1] and ProCancer-I[2] European AI4HI projects. No participants were able to trace back the identity of the patients, demonstrating the effectiveness of these pseudonymization profiles. Despite minor vulnerabilities identified in areas such as free-text metadata, 3D reconstructions, and usability limitations, no successful re-identification occurred—even with monetary rewards offered to participants.

Introduction

In the European Union, research projects handling personal data must adhere to the General Data Protection Regulation[3] (GDPR) to ensure the privacy of the individuals involved. Particularly, the GDPR Recital 26[4] defines anonymized data as "data rendered anonymous in such a manner that the data subject is not or no longer identifiable". In parallel, pseudonymization is defined by Article 4(5)[5] of the GDPR as "the processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organizational measures to ensure that the personal data are not attributed to an identified or identifiable natural person".

The secondary use of medical imaging datasets is a cornerstone for health imaging AI research. While the pseudonymization of the medical image metadata – the DICOM tags – is key, there are currently no consensus guidelines or standardized procedures to this end. Although some recommendations have been suggested by organizations like the NEMA[6], regulations leave the risk to data protection to the entity responsible for processing the data. Considering that the pseudonymization process requires the removal or replacement of any sensitive data that could potentially link to an individual, assessing the risk of re-identification through indirect identifiers is indeed challenging.

[1]Grupo de Investigación Biomédica en Imagen (GIBI230), Valencia, Spain. [2]Centro Hospitalar Universitário de Santo António (CHUdSA), Porto, Portugal. [3]Candiolo Cancer Institute, FPO-IRCCS, Turin, Italy. [4]Universitat de València (UV), Valencia, Spain. [5]Instituto de Instrumentación para Imagen Molecular (I3M), Universitat Politècnica de València (UPV), Valencia, Spain. [6]Hospital Universitario y Politécnico La Fe (HULAFE), Valencia, Spain. ✉e-mail: rocio_catalan@iislafe.es

MODALITY	ACQUISITION SITE		
	HULAFE	CHUdSA	IRCC
CT	Head/brain (5) Thorax-abdomen-pelvis (1)	Head/brain (1) Whole body (2)	Head/brain (1) Thorax (1)
MR	Head/brain (9) Liver (1)	Head/brain (1) Breast (1)	Head/brain (1) Abdomen (1)
CR	Sternon (1) Ribs (1) Thorax (1) Spinal column (1)	Chest (1) Hand (1)	Hip (1) Thorax (1)
PET-CT	Whole body (1) Head/brain (1)	—	—
RF	—	Knee (1)	—
PET	—	—	Whole body (2)

Table 1. Re-identification Challenge Dataset[8] Description.

While guaranteeing patient's privacy is crucial, image pseudonymization can reduce the research value of the dataset. To minimize the loss of non-personal, yet valuable data for research purposes, it is of paramount importance to evaluate the robustness of different pseudonymization methods, hence finding the right balance depending on the downstream applications.

Considering the urgent need to validate the pseudonymization techniques applied to medical imaging data; the goal of the present study was to engage the medical imaging community in attempting the re-identification of health images. This experiment, titled "The Re-identification Challenge", was conducted within the framework of the Artificial Intelligence for Health Imaging projects network[7] (AI4HI). As such, the objective of this exercise is to serve as a technical audit assessing the robustness of DICOM image pseudonymization methods implemented in two European AI4HI projects, ProCancer-I and ChAImeleon.

Materials and Methods

The development of the Re-identification Challenge involved several essential components, which will be presented sequentially.

Dataset[8] description. The 38 DICOM studies that constitute the challenge dataset[8] were acquired in three European acquisition sites: Hospital Universitario y Politécnico La Fe, Spain; Centro Hospitalar Universitário de Santo António, Portugal; and Candiolo Cancer Institute FPO-IRCCS, Italy. These studies encompass different modalities and anatomical regions. As a result, the dataset[8] includes a broader spectrum of DICOM tags, since the associated metadata can vary significantly depending on these factors. A detailed description of the dataset[8] is provided in Table 1.

Pseudonymization methods. The preparation and collection of the dataset[8] relied on three key components: the pseudonymization profiles from ChAImeleon and ProCancer-I, a customizable pseudonymization application, and de-facing software.

Each project within the framework of European initiatives that encompass medical studies pseudonymizes its data following customized procedures to comply with GDPR and national regulations. The objectives of these customized pseudonymization processes are maximizing the utility of the data for the project and safeguarding personal data. However, defining the pseudonymization protocols is a time-consuming and challenging task, primarily because regulations provide little technical guidance on how to do so, and there is no definitive assertion regarding re-identification risk. Therefore, each project defined its pseudonymization procedures[9].

The Clinical Trial Processor[10] application (CTP app) was utilized as the pseudonymization tool for the challenge dataset[8]. CTP is a stand-alone program developed by the Radiological Society of North America (RSNA) within the scope of the Medical Imaging Resource Center (MIRC) project. This application is widely recognized for its use in the pseudonymization of patient data in clinical trials and research projects, serving as the backend tool for initiatives within the AI4HI network, such as Procancer-I, as well as for various research groups such as Erasmus MC[11]. CTP is highly customizable, allowing its configuration to be tailored to specific needs. To meet these requirements, it was adjusted to incorporate the custom pseudonymization profiles developed for ChAImeleon and ProCancer-I. Afterwards, experts responsible for each project validated the proper implementation of these protocols, to assure faithful replication of them. CTP was distributed to the acquisition sites, enabling them to pseudonymize their studies at their sites, before sending them to the coordinator centre, HULAFE.

Once the dataset[8] was pseudonymized using the aforementioned methodology, the *mri_reface*[12] software was applied to a subset of head/brain studies for de-facing. *mri_reface* utilizes an average brain image template for each supported contrast. It first performs a linear registration followed by a nonlinear registration with the template, excluding facial and ear regions from the nonlinear transformation. The template is then warped to match the input image, with only a linear transformation applied to the face and ear regions. To ensure consistency, the template is further adjusted through intensity matching, bias field correction, and noise matching. Finally, the facial region in the target image is replaced with the corresponding region from the transformed and adjusted template[13]. In a known dataset of 157 volunteers, *mri_reface* lowered the matching rate from 97% to 28–30%. However, it also has drawbacks for brain measurements. Therefore, this study aimed to assess whether

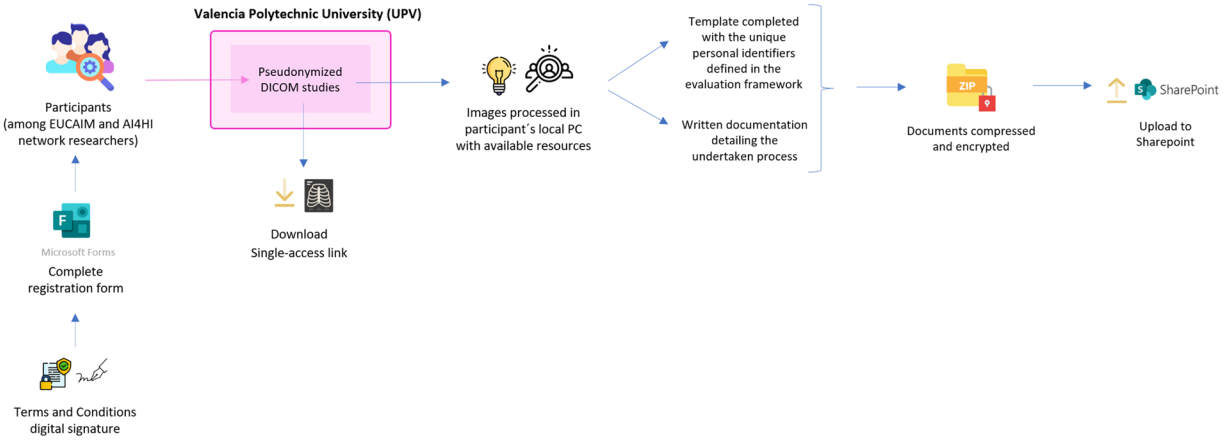

Fig. 1 Participant workflow in the Re-identification Challenge.

the re-identification risk in a broader, unrestricted setting—such as publicly available data—remained significant or if it was substantially reduced compared to having a closed dataset of photographs for comparison[12].

Participants registration and challenge dissemination. Regarding participants registration, a Microsoft Forms form was prepared, allowing individuals and groups to register. The primary requirement for registering in the challenge was to belong to AI4HI or EUCAIM projects partners. Additionally, participants were required to sign specific terms and conditions designed for the challenge. This ensured that access remained restricted, aligning with data protection and security compliance requirements.

The challenge dissemination was carried out through various channels, including project-related communication platforms, email lists of the researchers involved in the projects, and direct contact with their Principal Investigators (PIs) to ensure wide visibility among the target audience. As a result, the challenge reached a broad and relevant group of participants, most of whom participated in groups, which underscores the collaborative nature of this challenge.

Workflow and assessment methodology. During the execution of the challenge, participants relied on their own resources and technical infrastructure to carry out the tasks as they downloaded the dataset[8] hosted at Instituto de Instrumentación para Imagen Molecular (I3M). Over the course of three months, they focused on re-identifying the medical studies. Once their work was completed, they uploaded their results to a dedicated SharePoint space, ensuring an organized and efficient collection and review of their submissions. The requested documents included a filled template with the personal data to be evaluated, as well as a document where each participant or group explained their undertaken process. Additionally, each group was provided with a unique key to encrypt their results in a zip folder before uploading it to their SharePoint folder.

With respect to the evaluation methodology, it consisted of two distinct metric systems. The first involved assessing the re-identified patients, where re-identification was determined by successfully matching at least one unique personal identifier from the medical studies with the corresponding acquisition sites records. For this purpose, the following elements were considered unique identifiers: name and surnames; ID number (passport number, national identification number); phone number; full address; and e-mail address. These items were selected because they belong exclusively to a single individual and therefore constitute unequivocal proof of re-identification within our evaluation framework. If no participant succeeded in re-identifying any patients, a second evaluation system was introduced to assess the quality and rigor of the procedures followed in their re-identification attempts.

In addition, the concept of "singling out" from the GDPR is acknowledged, which defines a natural person as identifiable if it is reasonably likely that they could be singled out from a dataset, either directly or indirectly. In the context of this challenge, the dataset[8] provided to participants was fully pseudonymized and contained no direct identifiers. While participants could attempt re-identification by combining this information with external sources, a patient was considered successfully re-identified only if a unique personal identifier was obtained. This operational definition clarified that only information uniquely associated with an individual could be used to determine re-identification.

Figure 1 illustrates all the elements discussed, providing a visual representation from the participants' perspective, highlighting the process they had to undertake as part of the challenge.

Results

A total of 18 individuals and 11 teams, involving 68 participants in total, registered for the Re-identification Challenge. While no individual participants submitted results, 74% of the total participants submitted the requested documents. This high level of engagement highlights the collaborative spirit of the challenge and underscores the critical role of teamwork in addressing complex technical problems.

Despite the strong participation, none of the teams succeeded in tracing back the patients from the pseudonymized data. This outcome reinforces the robustness of the pseudonymization measures applied to the dataset[8], demonstrating their effectiveness in protecting patient privacy.

Nonetheless, the detailed reports provided by participants regarding their re-identification attempts were instrumental in uncovering minor vulnerabilities. While these vulnerabilities did not allow for direct re-identification, they did enable some teams to narrow down the potential population by identifying geographic regions associated with the data. However, it is important to note that this was only possible because there were three potential institutions, and participants could deduce them based on the language used in the metadata. These findings will be invaluable for further strengthening data pseudonymization techniques and ensuring even greater protection in future applications.

Seven teams submitted results, employing a wide variety of approaches. These included common strategies such as the automated extraction of DICOM metadata using the *pydicom*[14] Python library, followed by visual inspection to detect any identifiers present in the dataset[8]. As part of this inspection, some teams focused on specific attributes[15] more likely to contain sensitive information, such as these 50 attributes with potential to cause a security breach, either independently or in combination with other attributes.

Another frequently employed strategy involved inspecting the information contained in the images. On the one hand, they inspected whether any text was embedded within the images, either manually or using Optical Character Recognition (OCR) techniques and if so, identify any sensitive information that could directly or indirectly lead to re-identification. The applied OCR techniques included EasyOCR[16], Presidio Image Redactor[17] and Tesseract[18] via pytesseract[19]. On the other hand, with a special focus on head/brain studies, they attempted to detect facial and anatomical features such as scars and dental implants in the reconstructed 3D images to assess their potential for re-identification. Although the *mri_reface* software was applied to some studies, no successful identifications were made, highlighting the difficulty of re-identification even without face smoothing.

Several other approaches were used in the attempt to re-identify patients. For instance, one team cross-referenced the attributes *PatientSex* and *StudyDescription* to identify inconsistencies that might indicate modifications introduced by the pseudonymization process, aiming to infer the methods applied to the medical studies. Other teams analysed the language in *StudyDescription* and *SeriesDescription* tags and knowing the participating centres in the challenge, they were able to point out the acquisition sites of 21 studies, without further discoveries. Another attempt included having examined an implementation-specific use 128-byte section[20], which should be set to zeros if unused. In the challenge dataset[8], they reported that the field was indeed zeroed out, posing no risk of patient identification.

Some additional ideas, which were ultimately not implemented, were also discussed. One team proposed correlating the information stored in DICOM tags with instances of the project's XNAT databases, although they did not provide references of them. Another team, after conducting research regarding CTP app, put forward the idea of simulating synthetic data for some sensitive tags, apply MD5 hashing, embedding and autoencoder techniques and, eventually, develop machine learning models to extract complex patterns from the embeddings. Another idea that came up was to train a local AI model to cross-reference the reconstructed 3D head/brain images with internet sources although it was deemed unfeasible within its time and resource limits.

Moreover, some teams explored searching and correlating information from the internet. They reported gathering details about the projects involved as well as the pseudonymization tools used. Some teams conducted manual searches across public images related to the acquisition sites, and cancer patient associations but found no conclusive matches to the reconstructed 3D head/brain images. However, two teams were disqualified for violating the terms and conditions, which prohibited sharing challenge-related information online. One of them reported using OSINT techniques[21] that involved uploading screenshots of the reconstructed images. The other team leveraged various software tools, including *Pixel Enhancer*[22] from Canva to improve the resolution of the unsmoothed reconstructed images, reverse image search engines to automatically identify online matches, and *Age Verification*[23] application from *facia.ai*.

To provide a clearer overview of the techniques described in this section, Table 2 presents a summary of the approaches used or proposed by participants.

In addition, Table 3 presents the results of the second evaluation assessment, which qualitatively evaluates participants' re-identification approaches in terms of identification of vulnerabilities, creativity in proposed attack vectors, and the use of supporting references.

Discussion

The absence of successful re-identifications confirms the robustness of the pseudonymization techniques implemented, demonstrating their effectiveness in preventing linkage to patient identities. However, the analysis of participants' re-identification attempts revealed certain vulnerabilities that, while not leading to a complete breach, could pose potential risks under different conditions. In particular, issues related to free-text fields, reconstructed 3D images and usability limitations highlight the need for further refinement of pseudonymization strategies. Addressing these weaknesses is crucial to ensure long-term data protection while maintaining data usability, especially as re-identification techniques continue to evolve.

Free-text attributes. Participants' findings raised concerns related to free-text attributes in metadata, which may inadvertently contain sensitive information. Unexpected timestamps related to image acquisition were found in *SeriesDescription, IssueDateOfImagingServiceRequest*, and *ExposureDoseSequence* attributes. They also reported additional information such as clinical trial IDs in *PerformedProcedureStepDescription* and *ProtocolName* tags. Additionally, combining the information given in the challenge about the acquisition sites and the languages of the free-text in fields like *StudyDescription* and *SeriesDescription*, participants were able to infer the origin of most of the studies. While this information did not lead to re-identification, their presence emphasizes the risks posed

CATEGORY	SUMMARY
Widely adopted techniques	Automated DICOM metadata extraction using pydicom; manual inspection of tags; OCR using Tesseract, EasyOCR, and Presidio Image Redactor
Techniques used by a few teams	Detection of text embedded in images; analysis of facial/anatomical features in 3D reconstructions
Isolated technical explorations	Cross-referencing *PatientSex* and *StudyDescription* to detect inconsistencies; language-based analysis of tag content; checking 128-byte DICOM section
Proposed but unimplemented ideas	Correlating DICOM metadata with project XNAT instances; generating synthetic data; applying MD5 hashing and embeddings; training AI for 3D image matching
Disqualified due to rule violations	Uploading images online using various tools, including OSINT methods, Pixel Enhancer (Canva), reverse image search engines, and Age Verification (facia.ai)

Table 2. Summary of techniques used or proposed by participants.

	Vulnerabilities (40%)	Creativity (50%)	Citations and references (10%)	
	Participants analyze any weaknesses or failures detected in the pseudonymization process and propose potential mitigations	**Participants demonstrate innovative and creative solutions to discover new attack vectors in re-identification**	**Participants support their claims with reliable sources, academic citations, or reference materials that back up their methodology**	**Average**
Team 1	0	1	0	0.5
Team 2	6	1	3	3.2
Team 3	—	—	—	Disqualified
Team 4	1	1	2	1.1
Team 5	—	—	—	Disqualified
Team 6	0	1	0	0.5
Team 7	0	1	2	0.7

Table 3. Qualitative assessment of participants' re-identification approaches based on the second evaluation rubric.

by unstructured metadata. Incorporating natural language processing techniques, such as large language models (LLMs), into the pseudonymization process could enhance the detection of sensitive information in free-text fields. By systematically analysing these fields, such models could help identify and remove any inadvertently direct and indirect identifiers, reducing the risk of unintentional disclosure.

Reconstructed 3D head/brain studies. Another issue arising from the participants' findings concerns the matching of reconstructed 3D head/brain images that had not been de-faced. A previous study demonstrated that, given a known dataset of 157 volunteers, automated face recognition correctly matched 97% of MR medical images to photographs[12]. However, this challenge reflects a real-world scenario were participants lacked a reference database for comparison. Neither the manual efforts to cross-reference these medical studies with online databases nor the attempts by disqualified teams to use online software enabled a definitive match. The fact that no team, even with the motivation of monetary prizes, managed to identify the identity of the patients reinforces the effectiveness of the applied pseudonymization strategies.

Usability limitations derived from the pseudonymization profiles. Some participants raised concerns about possible inconsistencies in the pseudonymization profiles, particularly when opening studies with specific software such as *3D Slicer* and *ParaView*. They suggested that the performance of these profiles might lead to deviations from the DICOM standard. To assess these claims, the evaluation team reviewed the dataset[8] using *3D Slicer* and *OHIF viewer*.

In the ChAImeleon profile, every series opened correctly in both viewers. The only warnings encountered—such as attempts to load dynamic PET series as scalar volumes or doubled pixel spacing in processed images—were attributed to software-related limitations, specifically the 3D Slicer PET extension in the first case, and missing slices—originating from the dataset[8] itself rather than from a software limitation—in the second. These warnings also appeared in the original dataset[8], indicating that they were not due to pseudonymization issues.

In contrast, the ProCancer-I profile exhibited missing slices in dynamic diffusion and PET sequences after pseudonymization, compared to the original versions. This issue was traced to a DICOM tag not correctly handled during pseudonymization. Although these losses did not compromise patient privacy, they could impact the analysis of dynamic studies. This finding should be addressed in future protocol refinements.

Overall, ChAImeleon's profile was fully compliant with the DICOM standard. ProCancer-I also adhered to the standard, though issues related to slice loss were observed without compromising patient confidentiality. These findings highlight the need to refine pseudonymization protocols for better compatibility across processing tools and to preserve the usability of pseudonymized images.

To ensure GDPR compliance and protect patient privacy, participation in the challenge was restricted to specific project partners. Although this, along with the limited sample of DICOM studies, may reduce the diversity and generalizability of the results, such measures were necessary to minimize the risk of unauthorized access. Still, this approach marks a pioneering effort in re-identification research and sets a precedent for privacy-preserving methodologies in medical imaging.

Conclusion

This challenge sets an important precedent that may pave the way for robust evaluations of pseudonymization methods and may culminate in the standardization of pseudonymization profiles in the field of radiological imaging.

The robustness of the pseudonymization techniques was supported by the absence of successful re-identifications, even when participants explored vulnerabilities related to free-text fields, 3D head reconstructions, and usability limitations issues. Notably, despite the favorable context—limited number of known institutions and monetary prizes—no patient was re-identified, reinforcing the effectiveness of the methods used.

These findings underscore the importance of continuous refinement of pseudonymization protocols and provide valuable insights for future challenges in re-identification research, while also demonstrating the security and effectiveness of the methods developed in the ChAImeleon and ProCancer-I projects. As the field of radiological imaging moves toward standardized, privacy-preserving data sharing, this challenge sets a critical precedent for responsible research.

Data availability

The dataset used in this study is publicly available in Zenodo with https://doi.org/10.5281/zenodo.17034674[8].

Code availability

Several software tools were used during the Re-identification Challenge. The dataset[8] was pseudonymized using the CTP app (version released on 2023-11-04) with project-specific configurations. The CTP app is available in https://github.com/johnperry/CTP/raw/master/products/CTP-installer.jar and its documentation in https://mircwiki.rsna.org/index.php?title=MIRC_CTP. Moreover, the software used to apply de-facing in the dataset[8] is available in https://www.nitrc.org/projects/mri_reface.

Received: 10 September 2025; Accepted: 9 February 2026;
Published online: 08 April 2026

References

1. Bonmatí, L. M. *et al.* CHAIMELEON Project: Creation of a Pan-European Repository of Health Imaging Data for the Development of AI-Powered Cancer Management Tools. *Front. Oncol.* **12**, 742701. https://doi.org/10.3389/fonc.2022.742701 (2022).
2. ProCancer-I Project. An AI Platform integrating imaging data and health records for precision medicine in cancer management. https://www.procancer-i.eu/ (2025).
3. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation).
4. Regulation (EU) 2016/679, Recital 26.
5. Regulation (EU) 2016/679, Article 4.5.
6. NEMA is a contributor to the creation of DICOM standard and provides updated information on it. https://dicom.nema.org/medical/dicom/current/output/html/part15.html#chapter_E.
7. Kondylakis, H. *et al.* Data infrastructures for AI in medical imaging: a report on the experiences of five EU projects. *Eur Radiol Exp* **7**, 20, https://doi.org/10.1186/s41747-023-00336-x (2023).
8. Catalán Flores, R. *et al.* Re-identification Challenge Dataset [Data set]. *Zenodo* https://doi.org/10.5281/zenodo.17034674 (2025).
9. Kondylakis, H. *et al.* Documenting the de-identification process of clinical and imaging data for AI for health imaging projects. *Insights Imaging* **15**, 130, https://doi.org/10.1186/s13244-024-01711-x (2024).
10. RSNA MIRC. Clinical Trial Processor (CTP) documentation for clinical research and imaging data. https://mircwiki.rsna.org/index.php?title=MIRC_CTP (2025).
11. Erasmus MC University Medical Center. Research groups and facilities for medical imaging and health research. https://www.erasmusmc.nl/en/research/ (2025).
12. Schwarz, C. G. *et al.* Changing the face of neuroimaging research: Comparing a new MRI de-facing technique with popular alternatives. *NeuroImage* **231**, 117845, https://doi.org/10.1016/j.neuroimage.2021.117845 (2021).
13. Schwarz, C. G. *et al.* Effects of de-facing software mri_reface on utility of imaging biomarkers used in Alzheimer's disease research. *NeuroImage. Clinical* **40**, 103507, https://doi.org/10.1016/j.nicl.2023.103507 (2023).
14. Pydicom Contributors. Pydicom documentation: A pure Python package for working with DICOM files. https://pydicom.github.io/pydicom/stable/ (2025).
15. Aryanto, K. Y., Oudkerk, M. & van Ooijen, P. M. Free DICOM de-identification tools in clinical research: functioning and safety of patient privacy. *Eur Radiol* **25**, 3685–3695, https://doi.org/10.1007/s00330-015-3794-0 (2015).
16. JaidedAI. EasyOCR: Ready-to-use OCR with 80+ supported languages and all popular writing scripts. https://github.com/JaidedAI/EasyOCR (2025).
17. Microsoft. Microsoft Presidio: An open-source framework for detecting, redacting, and anonymizing sensitive data (PII). https://github.com/microsoft/presidio (2025).
18. Smith, R. An Overview of the Tesseract OCR Engine. *Ninth International Conference on Document Analysis and Recognition (ICDAR 2007)*, Curitiba, Brazil, pp. 629-633, https://doi.org/10.1109/ICDAR.2007.4376991 (2007).
19. Lee, M. A. pytesseract: A Python wrapper for Google Tesseract. GitHub repository, https://github.com/madmaze/pytesseract (2023).
20. NEMA. DICOM Part 10: Media Storage and File Format for Media Interchange. https://dicom.nema.org/medical/dicom/current/output/html/part10.html (2025).
21. OSINT Techniques. Resources for Open Source Intelligence and Social Media Investigations. https://osinttechniques.com/ (2025)..
22. Canva. Pixel Enhancer: AI-powered image resolution and quality enhancement tool. https://www.canva.com/apps/pixel-enhancer (2025).
23. Facia. Protect Children, Protect Your Business with Age Verification System. https://facia.ai/age-verification/ (2025).

Acknowledgements

Support from the ChAImeleon project (Horizon 2020, Grant Agreement No. 952172) is acknowledged for providing the framework and resources for organizing the Re-identification Challenge.

Author contributions

R.C.F., L.M.B., I.G.R.J., P.J.G. and R.M.M. contributed to the conception and design of the challenge. R.C.F., L.M.B. and I.G.R.J. contributed to writing the manuscript. R.C.F. was responsible for coordinating with the participating centres and participants. J.G.C., S.M., I.B., D.R., M.F. and L.M.B. contributed to the acquisition and preparation of the dataset[8]. J.D.S.Q. and P.L.L. prepared the secure environment where the dataset[8] was made available for participants to download. R.C.F., I.G.R.J., A.G.B. and L.M.B. reviewed the participants' results. RCF and P.J.G. were responsible for managing the challenge awards. All authors read, revised and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to R.C.F.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

© The Author(s) 2026