

On the Design of GDPR Compliant Workflows for Responsible Neuroimage Data Sharing*

Alexandros Karakasidis^{1,2}^a and Vassilios Vassalos¹

¹*Department of Informatics, Athens University of Economics & Business, Athens, Greece*

²*Department of Applied Informatics, University of Macedonia, Thessaloniki, Greece*

Keywords: Privacy Preservation, Data Sharing, Workflows, GDPR.

Abstract: Sharing medical data may facilitate advancing research as this may allow understanding the mechanisms of certain diseases, develop new drugs and medication schemes and find cures. However, as these data originate from humans, the issue of individual privacy rises since certain data modalities, as Neuroimages, if not properly curated, may reveal the identity of the individual described by these data. As legislation around the globe attempts to set rules for protecting privacy, techniques and methodologies have been proposed to allow for data publishing, also complying with the law. In this paper, we aspire to provide practitioners with workflows for ethical neuroimage data publishing under the GDPR, EU's latest data protection regulation.

1 INTRODUCTION

In the last years, medical research and operations are strongly relying on a variety of data modalities (Gkoulalas-Divanis and Loukides, 2015). In this paper, we focus on neuroimages, which are imaging data capturing the state or the operation of the human brain. Neuroimages exhibit explicit challenges and characteristics in terms of privacy preservation regarding data management and analysis. The reason is that neuroimage can allow identification either by revealing unique facial characteristics (Schimke and Hale, 2015) or, due to the uniqueness of the brain structure (Eke et al., 2021).

GDPR (European Union, 5 04) is EU's most recent legislation regarding storage, processing and management, including automated means, of personal data. However, GDPR does not explicitly define the measures required for privacy-preserving data sharing. Instead, it states that in terms of privacy preservation, the nature of processing, the state-of-the-art, the costs and the purposed should be taken into account. At this point, there are certain questions rising. First of all, what are the state-of-the-art methods for privacy-preserving neuroimage sharing? Then, how do they align with GDPR's requirements? Finally,


what happens regarding legal liability in the case of privacy breach?

Motivated by the fact that neuroimaging data are highly sensitive, we will try to answer these questions by first surveying state-of-the-art privacy preserving neuroimaging data sharing methods and by establishing three major categories of privacy protection mechanisms. Afterwards, based on this categorization, we design GDPR compliant workflows at the process level, to accommodate the reviewed privacy preservation methods, so as to investigate their compliance with GDPR provisions.

The rest of this paper is organized as follows. Section 2 presents works related to ours. In Section 3, there is the necessary background for our approach. Section 4 contains the mechanisms we propose and the categorization of state-of-the-art privacy-preserving sharing methods for neuroimages and Section 5 contains the proposed workflows based on these mechanisms. Finally, in Section 6 we sum up with our conclusions and present our next steps.

2 RELATED WORK

To the best of our knowledge, this is the first work attempting to describe GDPR compliant workflows for neuroimage sharing. Previous works in this area (Garjjo et al., 2014; Savio et al., 2017) as they are not

^a <https://orcid.org/0000-0001-7836-8444>

*Supported by the Human Brain Project (HBP): SGA3, Grant Agreement no 945539.

focused on privacy preservation they do not discuss GDPR implications either. In (Basin et al., 2018), Basin et al. propose a methodology for auditing GDPR compliance of business processes. However, this work does not regard sensitive personal information as in the case of neuroimages. In (Dumas et al., 2016), a method is proposed for analyzing differentially private workflows, however it is not evident how their methodology aligns with the GDPR or if it is applicable on neuroimage data. In (Besik and Freytag, 2019), authors use ontologies to check for privacy compliance in medical workflows, while in (Belhajjame et al., 2020), methods for identifying sensitive data in e-science data analysis workflows and their anonymization are presented. In these works, there is no distinction regarding different controllership setups. Finally, Besik et al. (Besik and Freytag, 2020) describe how consent should be managed under the GDPR, an approach that may be complementary to our approach in terms of implementation.

3 BACKGROUND

In this section, we provide the background knowledge required for laying out our methodology for sharing neuroimages under the GDPR. First, we briefly describe the key points of the GDPR for data processing and publishing. Then, we discuss some privacy-centered notions.

3.1 GDPR Key Points

The GDPR (European Union, 5 04) is the primary EU legislation concerning personal data processing and safeguarding individual privacy¹. It considers privacy by design, meaning that all data related operations should be designed so as to consider privacy at each of their steps, and privacy by default, which dictates that by default all available privacy measures should be applied each time personal data processing occurs. Furthermore, data should be maintained only as long as it is required for specific processing purposes.

Data processing, considers any action on data, e.g. storing, processing, analyzing and sharing. Without harming the general case, we consider a setup with two entities, the *data holder* and the *contractor*. The data holder owns the data, deciding how these data will be processed. Furthermore, the data holder may also perform some data processing as well. Not having the capacity for further data processing, the data

holder employees a contractor who will take up further data processing on behalf of the data holder.

To accommodate such situations, there is the definition of two roles. There is the *Data Controller*² and the *Data Processor*³. The data controller decides the purposes and means for data processing, while the data processor performs data processing on behalf of the data controller following her directions. In other words there is the capacity for outsourcing data processing, provided that the processor adheres to the controller's directions.

Of course, a data controller may process personal data as well. However, there is the case of joint controllership⁴, where a data processor also makes decisions with regard to processing means and purposes with some data controller. In this case, the legal liabilities change⁵. To this end, considering data sharing, the aforementioned involved entities should take all necessary measures so as for datasets containing personal information, to be responsibly processed, so that all these elements that make an individual identifiable to be removed. Next, we will briefly discuss the types of processing that personal data may undergo in order to achieve privacy preservation.

3.2 Anonymization, De-Identification and Pseudonymization

Anonymization, de-identification and pseudonymization are terms related with data processing, aiming at privacy and confidentiality preservation. Anonymization refers to the removal of private data from a record. In ISO 29100:2011 standard it is defined as the "process by which Personally Identifiable Information (PII) is irreversibly altered in such a way that a PII principal can no longer be identified directly or indirectly, either by the PII controller alone or in collaboration with any other party" (Eke et al., 2021).

In de-identification, a unique artificial code is assigned to data corresponding to each subject allowing its re-identification. Pseudonymization, according to GDPR, is very close to this description of de-identification (Eke et al., 2021). In pseudonymization, an individual may be identified using additional information. It is evident that these two processes are semantically close in the sense that in both cases there is some capacity for re-identification. Having these in mind, it is evident that all processes performed on personal data may not be identified as anonymization,

¹Article 7(1) of the GDPR

²Articles 4(7), 24 of the GDPR

³Articles 4(8), 28 of the GDPR

⁴Article 26 of the GDPR

⁵Recital 79 of the GDPR

unless they employ methods that aggregate data or add randomized noise to the data being processed, a fact however that may have negative impact on the respective utility of these data.

As a result, sharing of neuroimage data is not a trivial task and certain measures need to be taken. In (Eke et al., 2021), such safeguards and measures are described. These measures are: 1) informed consent, 2) data protection by design and by default and 3) data use agreements. As in this manuscript we are interested in technical measures that can be taken, we will focus on how data may be protected by design.

In order to comply with GDPR requirement for 'privacy by design' and 'privacy by default' for data sharing, applied measures comprise a combination of pseudonymization, encryption and access control. As we have earlier described pseudonymization, let us have a glimpse at the role of the other two measures that are required. Encryption, using secure hash functions is considered to meet current GDPR standards for data processing, data storage and data transfer, reducing the liability of a data controller (Eke et al., 2021). In terms of access control, now, since GDPR requires that, for privacy by design, "personal data are not made accessible to an indefinite number of natural persons or to bad actors" (Eke et al., 2021), using technical access control mechanisms emerges as a rational solution.

To sum up, data may be considered anonymous when they have been aggregated, minimized and distorted by random mechanisms. Otherwise, sharing should take place under the options that have been described, since such data may result to the individuals' re-identification.

4 MECHANISMS FOR SHARING NEUROIMAGE DATA

We will now present three mechanisms for categorizing privacy-preserving neuroimage data sharing depending on the privacy guarantees they provide and their release model. These mechanisms are generic and may be applicable to other data types as well. Then we categorize existing privacy-preserving sharing methods for neuroimages to each of these categories.

4.1 Data Sharing Mechanisms

Let us first begin by presenting our categorization, distinguishing three categories of data sharing mechanisms. As these mechanisms are generic, they may be also applicable beyond neuroimages. For each

of these mechanisms one or more privacy-preserving data publishing techniques may be used, each of them having its advantages and disadvantages. These are summarized in Table 1 and we will describe them in detail.

4.1.1 Non - Interactive Sharing with Formal Privacy Guarantees

The first mechanism we consider is a non-interactive release mechanism, which, however, will be able to provide formal privacy guarantees (NIF: Non - interactive Formal). We call this mechanism Non-Interactive as its based on the single release model as described earlier. For a method to provide formal privacy guarantees, a consistent mathematical model should exist. Furthermore, the privacy level offered may be determined through calculations and even calibrated. Differential privacy (Dwork, 2008) features these properties and is applicable to aggregate data (Kaissis et al., 2020; White et al., 2022). As such, the results of differentially private processing may be considered as fully anonymous, thus fulfilling the anonymization requirements GDPR sets.

On the other hand, there are certain drawbacks (Kaissis et al., 2020) when differential privacy is used. First of all, as differentially private mechanisms are applicable to aggregate data, a dataset of considerable size is required so as aggregation to be performed. Furthermore, since differential privacy based mechanisms are based on randomization, data are distorted. To this end, if accurate measurements are required, this may not be a suitable choice. Furthermore, differentially-private methods are parametric. This means that noise should be calibrated with caution not to undermine the utility of a dataset. Last but not least, such methods are not intuitive, requiring certain training in order to be used.

4.1.2 Non - Interactive Sharing with Non Formal Privacy Guarantees

The second release mechanism we consider also regards non-interactive data publishing but for methodologies that do not provide formal privacy guarantees (Non - Interactive, Non Formal: NINF). Such a method, like k-anonymity, may be used for a variety of reasons. For instance, such a method may already be included in some workflow, or there might even be no other alternative. Under these circumstances, such a method on its own may not be sufficient for providing privacy under the GDPR, thus requiring additional measures.

Access control mechanisms are categorized in this case as well. These do not pose formal privacy guar-

Table 1: Categorization of privacy-preserving data publishing techniques.

Mechanism	Sharing Technique	Advantages	Disadvantages
NIF	Differential Privacy (Dwork, 2008)	+ Formal Privacy Guarantees + Anonymization	- Datasets of considerable size required - Added noise distorts data - Needs careful calibration of amount of noise - Non intuitive
NINF	Observable characteristics removal: 1. Skull stripping (Kalavathi and Prasath, 2016) 2. Defacing (Schimke and Hale, 2011) 3. Face blurring (Milchenko and Marcus, 2013) 4. Refacing (Schwarz et al., 2021)	+ Intuitive + Maintains information (3)	- Reidentifiable - Requiring curation - Information Loss (1,2,4)
IF	Homomorphic Encryption (Acar et al., 2018)	+ Strong and provable security guarantees	- Computationally intensive - Does not protect identity / presence
	Secure Multiparty Computation (Zhao et al., 2019)	+ Strong and provable security guarantees	- Communication intensive - Does not protect identity / presence

antees in terms of an established mathematical model either. However, allowing authorized users only, this technique manages to block unauthorized access. Furthermore, since users have to provide credentials in order to access data, this allows for monitoring and tracking data usage. Last but not least, access control mechanisms are easy to implement, to use and to understand, as most practitioners are expected to have experience with user accounts. Regarding their other drawbacks, now, access control mechanisms are vulnerable to internal attacks in case of absence of additional privacy measures, since there is no protection of an individual's identity within a dataset with measures as generalization.

4.1.3 Interactive Sharing with Formal Privacy Guarantees

The third category of release mechanisms considers, this time, interactive data sharing. These are settings where computation is performed upon request and results are shared without releasing original data, while, at the same time, providing formal privacy guarantees (Interactive, Formal: IF). Such setups may utilize combinations of technologies as secure multiparty computation, federated learning or differential privacy in an interactive setting.

The question rising is whether this category of solutions adheres to the mandates of GDPR for data anonymization before publishing. The situation here is as follows. First of all, there is no direct publishing of data per se, but the results of computations. These computations should be secure, not allowing leakage of any personal information. This should be ensured through encryption mechanisms, as for instance Homomorphic Encryption, so as data transfers are secure. Similarly, in Secure Multiparty Computation techniques, data do not leave the participants.

Within these techniques all participating entities may assumed to be Data Processors, following the directions of the Data Controller.

4.2 Classes of Publishing Methods

Let us now see how existing neuroimage privacy-preserving sharing methods comply with the mechanisms we have described.

4.2.1 NIF

The first category of solutions, organized as NIF mechanisms, considers differentially private mechanisms (Kaissis et al., 2020). While this approach is suitable for all neuroimage types and able to provide formal privacy guarantees for the resulting dataset, certain issues arise (Sarwate et al., 2014) beyond those described earlier in Table 1. First, Neuroimaging data are often continuous-valued while differential privacy has mainly focused on discrete data. Next, neuroimaging datasets may contain few individuals and adding noise may significantly degrade utility. Furthermore, existing methods are application specific. This means that the entire dataset is not shared, but only specific metrics that will have to be decided in advance (e.g. the average value of some metric).

4.2.2 NINF

The second mechanism relates to well-known techniques for pseudonymizing neuroimages which alter the observable characteristics of a subject's face. These techniques are not sufficient to provide anonymity, thus their use should be complemented with organizational measures. In general, there are four directions in the literature. First there is face blurring or masking (e.g. (Milchenko and Marcus,

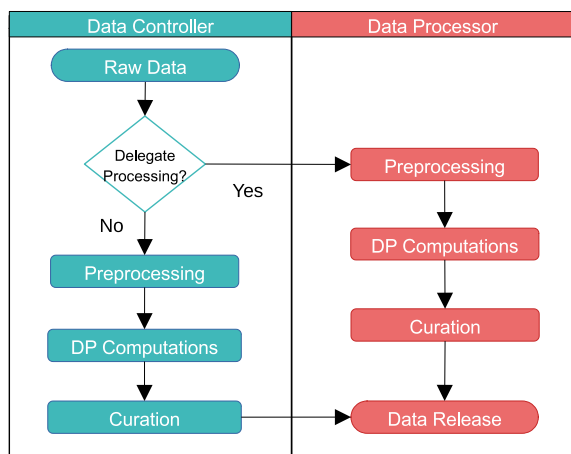


Figure 1: Distinct controllership NIF workflow.

2013)). Face blurring has been proven to be quite weak, allowing face reconstruction. Defacing cuts the entire face off from the neuroimage (Schimke and Hale, 2011), and finally, there is skull stripping. Skull stripping is the process of segmenting brain and non-brain elements to remove eyes, skin, etc and bone that may interfere with analyses (Schimke and Hale, 2015). Refacing replaces facial characteristics with averages (Schwarz et al., 2021).

As it is evident from their descriptions, these methods are only applicable to structural MRIs, other types of modalities such as fMRIs are not suitable for this publishing mechanism. Furthermore, defacing and refacing may allow some facial reconstruction based on skull contours. All these methods, however, fail to address linkage attacks, as the brain structure of the individual remains intact. Bearing these in mind, such data cannot be freely released in public. A solution towards using such a data modality would be to combine it with access control on the processor’s side and agreement signing for the users of these data (Eke et al., 2021).

4.2.3 IF

The third mechanism, as its name implies regards interactive solutions. Here, generic solutions based on Homomorphic Encryption and Secure Multiparty Computation may apply (Kaissis et al., 2020). Such solutions are the following. Neurocrypt (Senanayake et al., 2021) offers a suite based on Secure Multiparty Computation. Nevertheless, in this case as well, deployment was limited to less than ten participating sites. COINSTAC (Plis et al., 2016) is a dynamic decentralized platform. It can feature increased statistical power within its computations, since it may have access to multiple datasets, otherwise inaccessible. This framework is scalable, it supports a variety

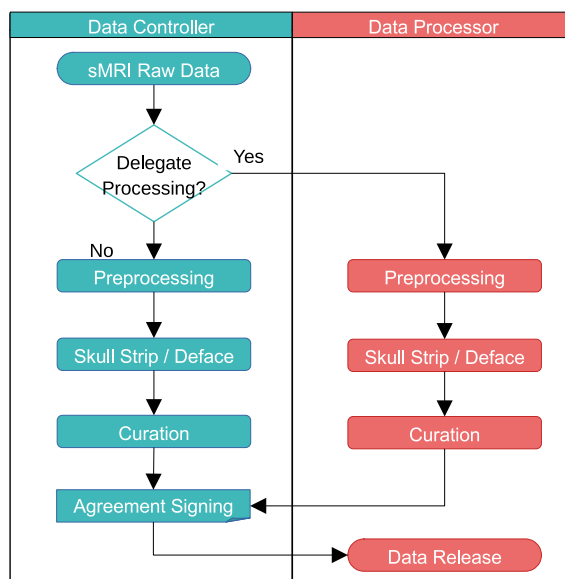


Figure 2: Distinct controllership NINF workflow.

of methods (Gazula et al., 2020) and has the capacity of enabling the incorporation of differentially private algorithms (Imtiaz and Sarwate, 2018). Further attempts in this area are in progress (Silva et al., 2020; Scherer et al., 2020). The Joint Imaging Platform⁶ of the German Cancer Consortium is established as a long term project for performing federated computation and analysis on distributed datasets. However, it is still in early stage of deployment. Fed-BioMed is a similar initiative by INRIA⁷, aligning with the aims of COINSTAC.

5 NEUROIMAGE PUBLISHING WORKFLOWS

We have seen how existing privacy-preserving publishing methods may be categorized to each mechanism. Now, for each of these mechanisms, we will provide workflows based on GDPR for the two controllership cases we have described, Distinct Controllership and Joint Controllership.

Data processing regards either computations on data or data sharing. Computations refer to any computational step towards changing the initial form of the data to enhance the privacy of the described individuals, e.g. defacing of a neuroimage, or even participation in a secure multiparty computation protocol. Such actions may be taken by the data processor, by the data controller, or by both. On the other hand, by

⁶<https://jip.dtk.dkfz.de/jiphompage/>

⁷<https://gitlab.inria.fr/fedbiomed>

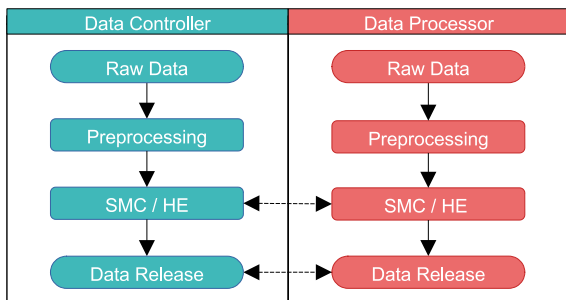


Figure 3: Distinct/joint controllership IF workflow.

data sharing, we indicate the action that makes data available beyond the owner.

At this point, it would be interesting to mention that data to be published may also require some *curation*, in the sense that there should be checks that processing has concluded as desired, that data is usable and that the type and amount of information shared is as planned. For instance, in case of MRI defacing, there should be checks that facial characteristics have been properly stripped off. Another example concerns the metadata of a neuroimage. Medical imaging formats as DICOM have the capacity to withhold identifying information. There should be proper measures for removing this information.

5.1 Distinct Controllership

Let us begin with the distinct controllership case. The controller, who also owns the data, decides the type of publishing she desires for her data, NIF, NINF, or IF, having taken into account the limitations and advantages of each of the aforementioned options. All decisions are taken at her side. The processor's role is to only perform processing, computation and sharing, according to controller's orders, thus being liable when processing deviates from controller's directions.

For NIF (Fig. 1), the controller is the one holding the raw data. She will have to decide whether she is performing data computations on her own, delegating only sharing to the processor. In this case, she performs all necessary preprocessing, computations (skull stripping or defacing), and curates the output. Now, the processed data, after differentially private computations take place, end up at the processor for release. Being differentially private, the may be assumed anonymized, thus freely shared. As the processor is only responsible for sharing, her implication with GDPR is limited to properly storing the data, without any loss or corruption, and for making them public with availability designated by the controller. Any curation is the controller's sheer responsibility, since the processor only hosts and shares the data.

Nevertheless, the controller may not have the tech-

nical experience or capacity to perform the required computations, delegating computations to the processor who has to process the data so as to comply with the requirements of the controller. Also, she will have to safeguard raw or preprocessed data for the period of time required and agreed and to perform curation on her side, since processing is her duty. Then, she shares the data as in the previous case. Someone may consider that the controller should also check for the curation results. However, we have assumed that there is no technical experience or capacity for her to do so, thus delegating these actions to the processor. Consequently, the processor is responsible for performing all these actions according to the controller's mandates. Of course, there is the possibility that the controller both processes and shares the data, however, we do not focus on this case as in this paper we study the implications of having two distinct cooperating entities.

For NINF processing, the workflow for is illustrated in Fig. 2. Here, we consider that the controller holds structural MRI data, suitable for defacing or skull stripping. This workflow resembles the one of the NIF case, but with some differences. First, processing only regards skull stripping or defacing. Second, and most important, there is an additional step of agreement signing for controlled release, as described in Section 3. As data of the NINF case are not considered to be anonymous, they cannot be freely released but they should follow a restricted access procedure. Thus, an agreement signing has to take place with the data controller who owns the data, so that processing purposes are agreed and access is tracked.

Finally, there is the third alternative with IF, which implies cooperated computation, releasing only results. In this case, all participants are performing computations. Data at the processor's side may originate from another controller not willing, or not having the capacity, to participate in the distributed computation. First of all, each of the participants should preprocess her data, to make them appropriate for the process, not revealing additional information. Next, each of the participants (controller and processor) process their local data as steps of multiparty computation schemes under the condition of using all available technical measures, as secure hash functions. This type of computations requires communication between the participating entities. Eventually, data release takes place as the successful calculation of the result of collaborative computations.

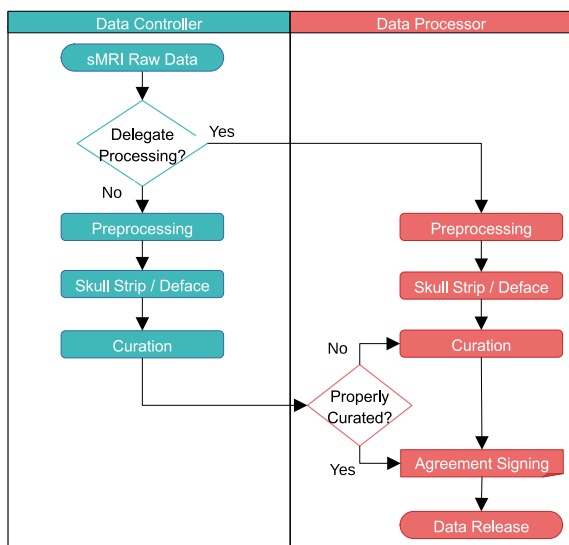


Figure 4: Joint controllership NINF workflow.

5.2 Joint Controllership

Joint Controllership’s case is more complicated as the processor also acts as a joint controller. There are additional GDPR implications on the processor’s side as well, as she also takes decisions regarding computations on data, having to ensure that the GDPR anonymization requirements are met.

Let us first visit the case of NIF release, as illustrated in Fig. 5. Again, there are two options. First, the data holder performs the differentially private calculations and the processor only shares the data. Second, both differentially private calculations and sharing are carried out by the processor. In the first case, the processor, also being liable under the GDPR as a joint controller, should also check that GDPR anonymization requirements have been covered by controller’s processing. To this end, the processor performs curation checking that personal identifiers have been excluded and that the differentially private method has been properly applied (e.g. proper method and parameters have been used so as to produce a properly anonymized dataset). If this has not been the case, she repeats curation also notifying the controller. In the second case, where the processor also performs calculations on data, she performs curation on the computation results by default.

For NINF release, as illustrated in Fig. 4, the same two cases as before apply. That is, data processing by the data holder and sharing by the processor, and both processing and sharing by the processor. Now, computation regards skull-stripping or defacing structural MRIs. As these data are not anonymous, the processor, as a joint controller should ensure that all nec-

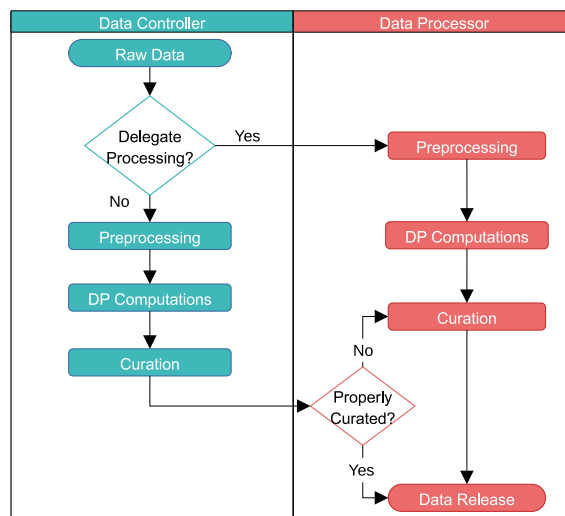


Figure 5: Joint controllership NIF workflow.

essary measures are taken to safeguard the subjects’ identities. As such, during curation, neuroimages are checked that personal identifiers have been removed and that defacing and skull stripping has been properly performed. If these are not the cases, data are accordingly processed. Then, these neuroimages are released through controlled access accompanied by an agreement signing so as to keep track of the users having access to these data. Agreement signing may be performed at processor’s side as she also has controller status due to joint controllership.

Finally, there is IF data sharing. As mentioned earlier, the data holder participates in a joint computation mechanism. As data remain at their owners, there is no need for curation and the workflow is identical with the distinct controllership case, as illustrated in Fig 3. Nevertheless, as the controllership is joint, the policies and details for this type of processing are decided by both parties.

6 CONCLUSIONS AND FUTURE WORK

In this paper, we have designed workflows for state-of-the-art privacy-preserving neuroimage data sharing techniques and discussed the implications for the main actors identified by the GDPR. Our next steps focus on two directions. First, we aim at investigating the implications of implementing the steps described at the process level and the respective mechanisms required for this purpose. Next, we are going to apply the lessons learned to design workflows for more data modalities (e.g. genomics).

ACKNOWLEDGEMENTS

Work performed while A. Karakasidis was cooperating with Athens University of Economics and Business. This research has been funded by the Human Brain Project (HBP) - SGA3, Grant agreement no 945539.

REFERENCES

- Acar, A., Aksu, H., Uluagac, A. S., and Conti, M. (2018). A survey on homomorphic encryption schemes: Theory and implementation. *ACM Computing Surveys*, 51(4):1–35.
- Basin, D., Debois, S., and Hildebrandt, T. (2018). On purpose and by necessity: compliance under the gdpr. In *International Conference on Financial Cryptography and Data Security*, pages 20–37. Springer.
- Belhajjame, K., Faci, N., Maamar, Z., Burégio, V., Soares, E., and Barhamgi, M. (2020). On privacy-aware escience workflows. *Computing*, 102(5):1171–1185.
- Besik, S. I. and Freytag, J.-C. (2019). Ontology-based privacy compliance checking for clinical workflows. In *LWDA*, pages 218–229.
- Besik, S. I. and Freytag, J.-C. (2020). Managing consent in workflows under gdpr. In *ZEUS*, pages 18–25.
- Dumas, M., García-Bañuelos, L., and Laud, P. (2016). Differential privacy analysis of data processing workflows. In *International Workshop on Graphical Models for Security*, pages 62–79. Springer.
- Dwork, C. (2008). Differential privacy: A survey of results. In *International conference on theory and applications of models of computation*, pages 1–19. Springer.
- Eke, D., Aasebø, I. E., Akintoye, S., Knight, W., Karakasidis, A., Mikulan, E., Ochang, P., Ogoh, G., Oostenveld, R., Pigorini, A., et al. (2021). Pseudonymisation of neuroimages and data protection: increasing access to data while retaining scientific utility. *Neuroimage: Reports*, 1(4):100053.
- European Union (2016-05-04). Regulation 2016/679. *Official Journal L110*, 59:1–78.
- Garijo, D., Corcho, O., Gil, Y., Braskie, M. N., Hibar, D., Hua, X., Jahanshad, N., Thompson, P., and Toga, A. W. (2014). Workflow reuse in practice: A study of neuroimaging pipeline users. In *2014 IEEE 10th International Conference on e-Science*, volume 1, pages 239–246. IEEE.
- Gazula, H., Kelly, R., Romero, J., Verner, E., Baker, B. T., Silva, R. F., Imtiaz, H., Saha, D. K., Raja, R., Turner, J. A., et al. (2020). Coinstac: Collaborative informatics and neuroimaging suite toolkit for anonymous computation. *Journal of Open Source Software*, 5(54):2166.
- Gkoulalas-Divanis, A. and Loukides, G. (2015). *Medical data privacy handbook*. Springer.
- Imtiaz, H. and Sarwate, A. D. (2018). Distributed differentially private algorithms for matrix and tensor factorization. *IEEE journal of selected topics in signal processing*, 12(6):1449–1464.
- Kaissis, G. A., Makowski, M. R., Rückert, D., and Braren, R. F. (2020). Secure, privacy-preserving and federated machine learning in medical imaging. *Nature Machine Intelligence*, 2(6):305–311.
- Kalavathi, P. and Prasath, V. (2016). Methods on skull stripping of mri head scan images—a review. *Journal of digital imaging*, 29(3):365–379.
- Milchenko, M. and Marcus, D. (2013). Obscuring surface anatomy in volumetric imaging data. *Neuroinformatics*, 11(1):65–75.
- Plis, S. M., Sarwate, A. D., Wood, D., Dieringer, C., Landis, D., Reed, C., Panta, S. R., Turner, J. A., Shoemaker, J. M., Carter, K. W., et al. (2016). Coinstac: a privacy enabled model and prototype for leveraging and processing decentralized brain imaging data. *Frontiers in neuroscience*, 10:365.
- Sarwate, A. D., Plis, S. M., Turner, J. A., Arbabshirani, M. R., and Calhoun, V. D. (2014). Sharing privacy-sensitive access to neuroimaging and genetics data: a review and preliminary validation. *Frontiers in neuroinformatics*, 8:35.
- Savio, A. M., Schutte, M., Graña, M., and Yakushev, I. (2017). Pypes: workflows for processing multimodal neuroimaging data. *Frontiers in Neuroinformatics*, 11:25.
- Scherer, J., Nolden, M., Kleesiek, J., Metzger, J., Kades, K., Schneider, V., Bach, M., Sedlaczek, O., Bucher, A. M., Vogl, T. J., et al. (2020). Joint imaging platform for federated clinical data analytics. *JCO Clinical Cancer Informatics*, 4:1027–1038.
- Schimke, N. and Hale, J. (2011). Quickshear defacing for neuroimages. In *HealthSec*.
- Schimke, N. and Hale, J. (2015). Privacy considerations and techniques for neuroimages. In *Medical Data Privacy Handbook*, pages 527–547. Springer.
- Schwarz, C. G., Kremers, W. K., Wiste, H. J., Gunter, J. L., Vemuri, P., Spsychalla, A. J., Kantarci, K., Schultz, A. P., Sperling, R. A., Knopman, D. S., et al. (2021). Changing the face of neuroimaging research: Comparing a new mri de-facing technique with popular alternatives. *NeuroImage*, 231:117845.
- Senanayake, N., Podschwadt, R., Takabi, D., Calhoun, V. D., and Plis, S. M. (2021). Neurocrypt: Machine learning over encrypted distributed neuroimaging data. *Neuroinformatics*, pages 1–18.
- Silva, S., Altmann, A., Gutman, B., and Lorenzi, M. (2020). Fed-biomed: A general open-source frontend framework for federated learning in healthcare. In *Domain Adaptation and Representation Transfer, and Distributed and Collaborative Learning*, pages 201–210. Springer.
- White, T., Blok, E., and Calhoun, V. D. (2022). Data sharing and privacy issues in neuroimaging research: Opportunities, obstacles, challenges, and monsters under the bed. *Human Brain Mapping*, 43(1):278–291.
- Zhao, C., Zhao, S., Zhao, M., Chen, Z., Gao, C.-Z., Li, H., and Tan, Y.-a. (2019). Secure multi-party computation: theory, practice and applications. *Information Sciences*, 476:357–372.