Privacy Preserving Federated Deep Learning Infrastructure for ARGOS: ARtificial intelligence for Gross tumor vOlume Segmentation.

Ananya Choudhury, Leory Volmer, Frank Martin, RRR Fijten, Leonard Wee, Andre Dekker, Johan van Soest

Submitted to: JMIR AI
on: May 23, 2024

Table of Contents

Privacy Preserving Federated Deep Learning Infrastructure for ARGOS: ARtificial intelligence for Gross tumor vOlume Segmentation.

Ananya Choudhury[1, 2*] MSc; Leory Volmer[1, 2*]; Frank Martin[3]; RRR Fijten[1, 2]; Leonard Wee[1, 2]; Andre Dekker[1, 2, 4]; Johan van Soest[1, 2, 4]

[1]GROW Research Institute for Oncology and Reproduction Maastricht University Medical Center+ Maastricht NL
[2]Clinical Data Science Maastricht University Maastricht NL
[3]Netherlands Comprehensive Cancer Organization (IKNL) Eindhoven NL
[4]Brightlands Institute for Smart Society (BISS) Faculty of Science and Engineering (FSE) Maastricht University Heerlen NL
[*]these authors contributed equally

Corresponding Author:
Ananya Choudhury MSc
GROW Research Institute for Oncology and Reproduction
Maastricht University Medical Center+
Paul Henri Spakalaan
1
Maastricht
NL

Abstract

Background: Accurate delineation of the gross tumor volume (GTV) is crucial in radiotherapy for dose calculation and precise imaging-guided treatment of lung cancer patients. Conventionally, this task has been performed manually by radiation oncologists, which can be subjective and vary among clinicians. Deep learning has enabled automated GTV segmentation, with the potential to revolutionize the radiotherapy workflow by improving efficiency and consistency, ultimately enhancing patient outcomes while reducing clinician workload. However, the adoption of deep learning based GTV segmentation tools is hindered by the challenges of data privacy and the need for large, diverse datasets across multiple institutions. Federated learning (FL) offers a promising solution, allowing collaborative development of AI models without the need to share individual subject-level data.

Objective: The objective is to introduce an innovative federated learning infrastructure called the Personal Health Train (PHT) that includes the procedural, technical, and governance components needed to implement federated learning on real-world healthcare data, including training deep learning neural networks. The study aims to apply this federated deep learning infrastructure to the use case of gross tumor volume (GTV) segmentation on chest CT images of lung cancer patients, and present the results from a proof-of-concept experiment.

Methods: The PHT framework addresses the challenges of data privacy concerns of data sharing by keeping data close to the source, and instead sending analysis to the data. Technologically, PHT requires three interdependent components: "tracks" (protected communication channels), "trains" (containerized software applications), and "stations" (institutional data repositories), which are supported by the open source "Vantage6" software. The study applies this federated deep learning infrastructure to the use case of GTV segmentation on chest CT images of lung cancer patients, with the introduction of an additional component called the Secure Aggregations Server, where the model averaging is done in a trusted and inaccessible environment.

Results: In this paper we demonstrated the feasibility of executing deep learning algorithms in a federated manner using PHT and presented the results from a proof-of-concept study. The infrastructure linked 12 hospitals across 8 nations, covering 4 continents, demonstrating the scalability and global reach of the proposed approach. In the entire execution and training of the deep learning algorithm, no data has been shared outside the hospital.

Conclusions: The findings of the proof-of-concept study, as well as the implications and limitations of the infrastructure and the results, are discussed. The application of federated deep learning to unstructured medical imaging data, facilitated by the PHT framework and Vantage6 platform, represents a significant advancement in the field. The proposed infrastructure addresses the

challenges of data privacy and enables collaborative model development, paving the way for the widespread adoption of deep learning-based tools in the medical domain and beyond. The introduction of the Secure Aggregation Server implied that data leakage problems in federated learning can be prevented by careful design decisions of the infrastructure. Clinical Trial: ARtificial Intelligence for Gross Tumour vOlume Segmentation (ARGOS)
ClinicalTrials.gov ID NCT05775068
Sponsor Maastricht Radiation Oncology
Information provided by Andre Dekker, Maastricht Radiation Oncology (Responsible Party)
https://clinicaltrials.gov/study/NCT05775068

(JMIR Preprints 23/05/2024:60847)
DOI: https://doi.org/10.2196/preprints.60847

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✓ **Please make my preprint PDF available to anyone at any time (recommended).**
Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
Only make the preprint title and abstract visible.
No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**
Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v
Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

Original Manuscript

Privacy Preserving Federated Deep Learning Infrastructure for ARGOS: ARtificial intelligence for Gross tumor vOlume Segmentation.

Abstract

Background

Accurate delineation of the gross tumor volume (GTV) is crucial in radiotherapy for dose calculation and precise imaging-guided treatment of lung cancer patients. Conventionally, this task has been performed manually by radiation oncologists, which can be subjective and vary among clinicians. Deep learning has enabled automated GTV segmentation, with the potential to revolutionize the radiotherapy workflow by improving efficiency and consistency, ultimately enhancing patient outcomes while reducing clinician workload. However, the adoption of deep learning based GTV segmentation tools is hindered by the challenges of data privacy and the need for large, diverse datasets across multiple institutions. Federated learning (FL) offers a promising solution, allowing collaborative development of AI models without the need to share individual subject-level data.

Objective

The objective is to introduce an innovative federated learning infrastructure called the Personal Health Train (PHT) that includes the procedural, technical, and governance components needed to implement federated learning on real-world healthcare data, including training deep learning neural networks. The study aims to apply this federated deep learning infrastructure to the use case of gross tumor volume (GTV) segmentation on chest CT images of lung cancer patients, and present the results from a proof-of-concept experiment.

Methods

The PHT framework addresses the challenges of data privacy concerns of data sharing by keeping data close to the source, and instead sending analysis to the data. Technologically, PHT requires three interdependent components: "tracks" (protected communication channels), "trains" (containerized software applications), and "stations" (institutional data repositories), which are supported by the open source "Vantage6" software. The study applies this federated deep learning infrastructure to the use case of GTV segmentation on chest CT images of lung cancer patients, with the introduction of an additional component called the Secure Aggregations Server, where the model averaging is done in a trusted and inaccessible environment.

Results

In this paper we demonstrated the feasibility of executing deep learning algorithms in a federated manner using PHT and presented the results from a proof-of-concept study. The infrastructure linked 12 hospitals across 8 nations, covering 4 continents, demonstrating the scalability and global reach of the proposed approach. In the entire execution and training of the deep learning algorithm, no data has been shared outside the hospital.

Discussion and Conclusion

The findings of the proof-of-concept study, as well as the implications and limitations of the infrastructure and the results, are discussed. The application of federated deep learning to unstructured medical imaging data, facilitated by the PHT framework and Vantage6 platform, represents a significant advancement in the field. The proposed infrastructure addresses the challenges of data privacy and enables collaborative model development, paving the way for the widespread adoption of deep learning-based tools in the medical domain and beyond. The introduction of the Secure Aggregation Server implied that data leakage problems in federated learning can be prevented by careful design decisions of the infrastructure.

Introduction

Federated Learning (FL) allows collaborative development of artificial intelligence models using large datasets, without the need to share individual subject-level data [1, 2, 3, 4]. In FL, partial models trained on separate datasets are shared, but not the data itself, hence a global model is derived from the collective set of partial models. This study introduces an innovative federated learning framework known as the Personal Health Train (PHT) that includes the procedural, technical and governance needed to implement FL on real world healthcare data, including the training of deep learning neural networks[5]. The PHT infrastructure is supported by a free and open-source infrastructure known as “priVAcy preserviNg federaTed leArninG infrastructurE for Secure Insight eXchange” i.e. Vantage6[6]. We will describe in detail an architecture for training a deep-learning model in a federated way with twelve institutional partners located in different parts of the world.

Sharing patient data between healthcare institutions is tightly regulated due to concerns about patient confidentiality and the potential for misuse of data. Data protection laws – including the European Union's General Data Protection Regulations (GDPR), Health Insurance Portability and Accountability Act of 1996 (HIPAA) in the United States of America, and similar regulations in China, India, Brazil, and many other countries - place strict conditions on the sharing and secondary use of patient data[7]. Incompatibilities between laws and variation in the interpretation of such laws leads to strong reluctance about sharing data across organizational and jurisdictional boundaries[8,9,10].

To address the challenges of data privacy, a range of approaches have been published in the literature. Differential privacy, homomorphic encryption, and federated learning comprise a family of applications known as “privacy enhancing technologies” (PETs)[11,12,13,]. The common goal of PETs is to unlock positively impactful societal, economic, and clinical knowledge by analyzing data en masse, while obscuring the identity of study subjects that make up the dataset. Academic institutions are more frequently setting up controlled workspaces (e.g. secure research environments, a.k.a. SREs) where multiple researchers can collaborate on data analysis within a common cloud computing environment, but without allowing access to the data from outside the SRE desktop, however this assumes that all the data needed has been transferred into the SRE in the first place[14,15]. Similarly, the National Institutes of Health (NIH) has set up an "Imaging Data Commons" to provide secure access to a large collection of publicly available cancer imaging data co-located with analysis tools and resources[16]. Other researchers have shown that blockchain encryption technology can be used to securely store and share sensitive medical data[17]. Blockchain ensures data integrity by maintaining an audit trail of every transaction while zero trust principles make sure the medical data is encrypted and only authenticated users and devices interact with the network[18].

From a procedural point of view, the PHT manifesto for FL rules out sharing of individual patient-level data between institutions, no matter if the patient data has been de-identified or encrypted [19]. The privacy-by-design principle here may

be referred to as “safety in numbers”, i.e. any singles individual’s data values are obscured, by computing either the descriptive statistics or the partial model, over multiple patients. PHT allows sufficiently adaptable methods of model training, such as iterative numerical approximation (e.g. bisection) or federated averaging (FedAvg[20]), and does not mandatorily require model gradients or model residuals, which are well-known avenues of privacy attacks[21,22,23,24]. Governance is essential with regards to compliance with privacy legislation, and division of intellectual property between collaboration partners. A consortium agreement template for PHT has been made openly accessible[25], that is based on our current consortium ARGOS (ARtificial intelligence for Gross tumor vOlume Segmentation)[26] . Technologically, PHT requires 3 interdependent components to be installed—“tracks” are protected telecommunications channels that connect partner institutions, “trains” are Docker containerized software applications that execute a statistical analysis which all partners have agreed upon, and “stations” are the institutional data repositories which hold the patient data[23]. It is this technological infrastructure – the tracks, trains and stations – that are supported by the aforementioned Vantage6 software, for which detailed standalone documentation exists[27].

Recently, the application of deep learning in healthcare has led to impressive results, specifically in the areas of natural language processing and computer vision (medical image analysis), with the promise for more efficient diagnostics and better predictions of treatment outcomes in future[28,29,30,31,32] . However, for robust generalizability, and to earn clinicians’ acceptance, it is essential that AI applications are trained on massive volumes of diverse and demographically representative healthcare data across multiple institutions. Given the barriers to data sharing, this is clearly an area where FL can play a vital role. Many studies have been published which presents federated learning on medical data including federated deep learning[33,34,35,36,37]. However, only a limited number of studies have documented the use of dedicated frameworks and infrastructures in a transparent manner. Adoption of custom federation strategy or absence of explicit reporting on the utilized infrastructure is observed in most of the studies. Table 1 summarizes the small number of federated learning studies that have been published in connection with deep learning investigations related to medical image segmentations to date.

Table 1: Existing studies from the literature focussing on federated deep learning on medical images

Infrastructure	Clinical Use Case	Data Type	Scale
NVIDIA FLARE/ CLARA	Prostate Segmentation of T2-weighted MRI[38]	DICOM MRI	3 centers
	COVID-19 Pneumonia Detection[39]	Chest CT	7 centers
Tensorflow Federated	COVID-19 prediction from chest CT Images[40]	Chest CT	3 datasets
OpenFL	Glioblastoma Tumor Boundary Detection[41]	Brain MRI	71 centers

The use case we have selected for this work is gross tumor volume (GTV) segmentation on chest computed tomography (CT) images of persons diagnosed with lung cancer. An example is shown in Figure 1. Accurate delineation of the GTV is crucial in radiotherapy for dose calculation and precise imaging-guided treatment. Conventionally, this work has been done manually

by highly qualified radiation oncologists, with potential for subjectivity and variation among clinicians. Deep learning has made feasible automated GTV segmentation Adoption of these deep learning based GTV segmentation tools has the potential to revolutionize the radiotherapy workflow, improving the efficiency and consistency of the treatment planning process, and ultimately enhancing patient outcomes while reducing clinician workload.

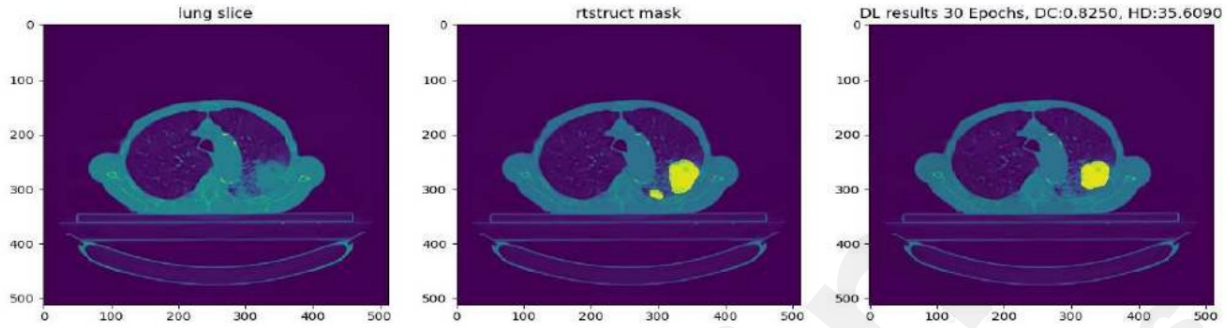

Figure 1: Illustrative result on a hold-out validation slice; the main bulk of the GTV as determined by the oncologist (middle) has been correctly delineated by the deep learning algorithm (right), but a small tumor mass adjacent and to the lower right of the main GTV mass has been missed. Reproduced from Figure 6 of Chapter 4 of the Patil thesis[42].

The paper proposes a federated deep learning infrastructure based on the Personal Health Train (PHT) Manifesto[19], which provides a governance and ELSI (Ethical, Legal, and Social Implications) framework for conducting federated learning studies across geographically diverse data providers. We explore applying federated learning to unstructured medical imaging data, which is the first of its kind, and develop a custom infrastructure using the open-source Vantage6 platform for making it suitable for deep learning-based studies. The subsequent sections provide a comprehensive account of the precise technical specifications of the infrastructure that links 12 hospitals across 8 nations, covering 5 continents. We developed a Convolutional Neural Network based algorithm to learn from the distributed datasets and deploy it using the infrastructure. The findings of the proof-of-concept study as well as the implications and limitations of the infrastructure and the results, are discussed. The paper is structured as follows: the methods section describes the approach taken, followed by the results which detail the implementation of the infrastructure and a proof-of-concept execution. Finally, the paper concludes with a discussion and conclusion section.

Methods

When conducting a federated deep learning study, it is crucial to consider several key perspectives which include both technical as well as organizational and legal aspects. These key factors have been instrumental in designing the infrastructure architecture used for training the deep learning algorithm. In this section we discuss the technical details while adhering to an Ethics-Legal-Social Impact framework as laid down by the PHT manifesto. The technical design decisions are based on the following assumptions:

Data landscape: Understanding the data landscape is crucial in designing and deploying federated learning algorithms. The technological approaches for

handling horizontally partitioned data, where each institution contains non-overlapping human subjects but the domain of the data (e.g. CT images of lung cancer) is the same across different institutions, can differ significantly from those used for vertically partitioned data, where each institution contains the same human subjects but the domain of the data do not overlap (e.g. CT scans in one, but socio-economic metrics in another). Additionally, unstructured data, such as medical images, requires different algorithms and preprocessing techniques compared to structured data. In this paper, the architecture will only focus on CT scans and horizontally partitioned patient data.

Data Preprocessing: In a horizontally partitioned federated learning setting, the key preprocessing steps can be standardized and sent to all partner institutions. However, the workflow needs to handle differences in patients, scan settings, and orientations. Anonymization, quality improvements, and DICOM standardization ensure homogeneity and high quality across hospitals. These offline preprocessing steps, applied consistently to the horizontally partitioned data, enabled using the same model across institutions, crucial for the federated learning study's success.

Network topology of the FL Infrastructure: The network topology choice for implementing FL can vary from client-server, peer-to-peer, tree based hierarchical or hybrid topologies. While peer-to-peer architecture is more cost-effective and offers a high capacity, it has the disadvantages of a lack of security and privacy constraints and a complex troubleshooting process in the event of a failure. The choice of network topology for this study is based on a client-server architecture, offering a single point of control in the form of the central server.

Choice of Model Aggregation Site: For a client-server architecture, the model aggregation can occur either in one of the data providers machines, or the central server or in a dedicated aggregation server. For this implementation, we opted for using a dedicated aggregation server. The details and benefits of the implementation are discussed in the next section.

Training Strategy: The communication mechanism for transferring weights can be either synchronous, asynchronous, or semi-synchronous and weights can be consolidated using ensemble learning, federated averaging, split learning, weight transfer or swarm learning. The strategy employed for this study is based on a synchronous mechanism using the federated averaging (FedAvg) algorithm. This gives a simple approach, where the averaging algorithm waits for all the data centers to transfer the locally trained model before initiating the averaging.

Based on the assumption, Figure 2 depicts an overall architecture of the federated deep learning study presented in the paper. The next section describes the FL Infrastructure in detail.

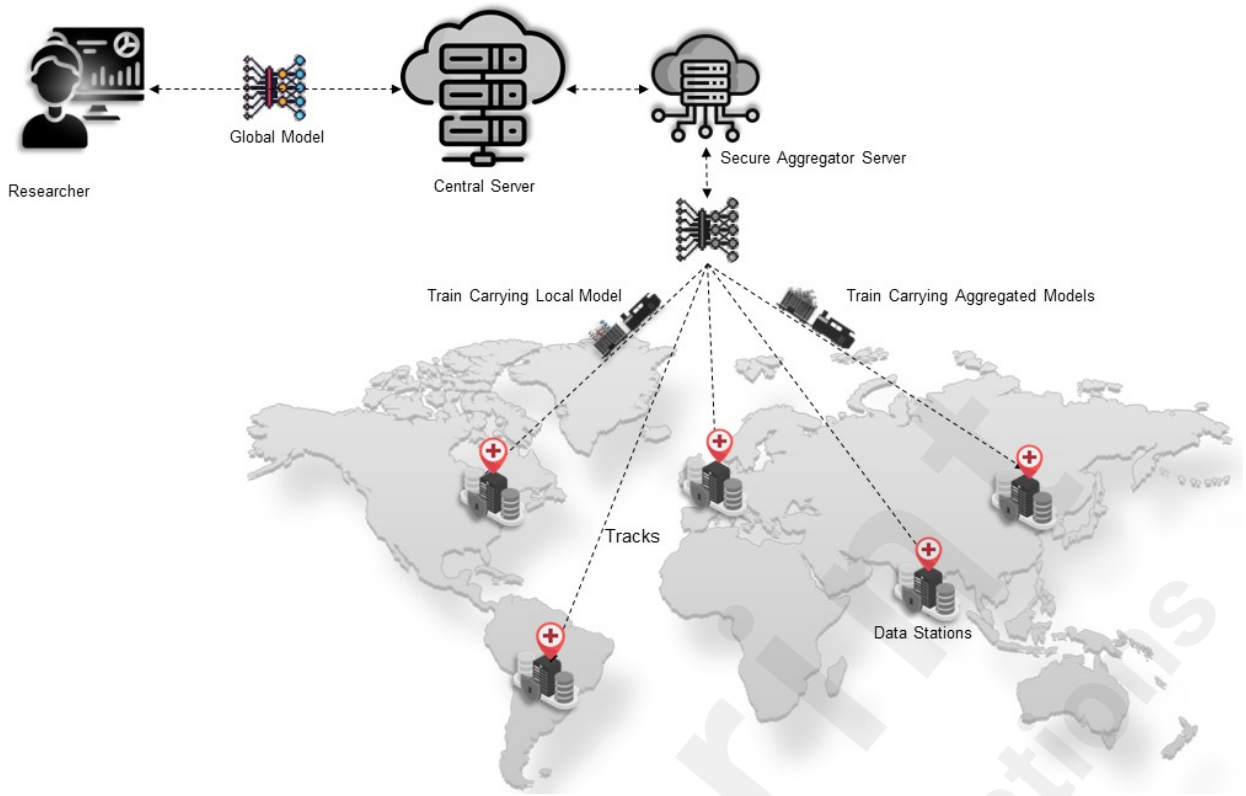

Figure 2: Overall Architecture of ARGOS Federated Deep Learning architecture adapted from Vantage6. The figure depicts a researcher connected to the central server, a secure aggregation server, trains carrying models, connected data stations and the communicating tracks.

The ARGOS Federated Deep Learning Infrastructure

In accordance with the PHT principles, the ARGOS infrastructure is comprised of three primary categories of components, labeled as the Data Stations, the Trains, and the Track. Furthermore, the architectural framework encompasses various roles which map to the level of permissions and access, specifically a track provider, the data providers, and a researcher. The infrastructure implementation can be further categorized into three important components: a central coordination server, a Secure Aggregation Server (SAS), and the nodes located at each 'Data Stations'. Following we attempt to describe each of these components and the respective stakeholders responsible for maintaining them.

Central Coordinating Server

The central coordination server is located at the highest hierarchical level and serves as an intermediary for message exchange among all other components. The components of the system, including the users, data stations, and secure aggregation server, are registered entities that possess well-defined authentication mechanisms within the central server. It’s noteworthy that the central acts as a coordinator rather than a computational engine. Its primary function is to store task-specific metadata relevant to the task initiated for training the deep learning algorithm. In the original Vantage6 infrastructure, the central server also stores the intermediate results. In the ARGOS infrastructure, the central server is designed to not store any intermediate results but only the global aggregated model at the end of the entire training process.

Secure Aggregation Server

The SAS refers to a specialized station that contains no data and functions as a consolidator of locally trained models. The aggregator node is specifically designed to possess a Representational State Transfer (REST) – Application Programming Interface (API) termed as the *API Forwarder*. The *API Forwarder* is responsible for managing the requests received from the Data Stations and subsequently routing them to the corresponding active Docker container, running the aggregation algorithm.

To prevent any malicious or unauthorized communication with the aggregator node, each data station is equipped with a JSON Web Token (JWT) that is unique for each iteration. The *API Forwarder* only accepts communications that are accompanied by a valid JWT. The implementation of this functionality guarantees the protection of infrastructure users and effectively mitigates the risk of unauthorized access to the SAS. Figure 3 shows the architecture and execution mechanism for the SAS.

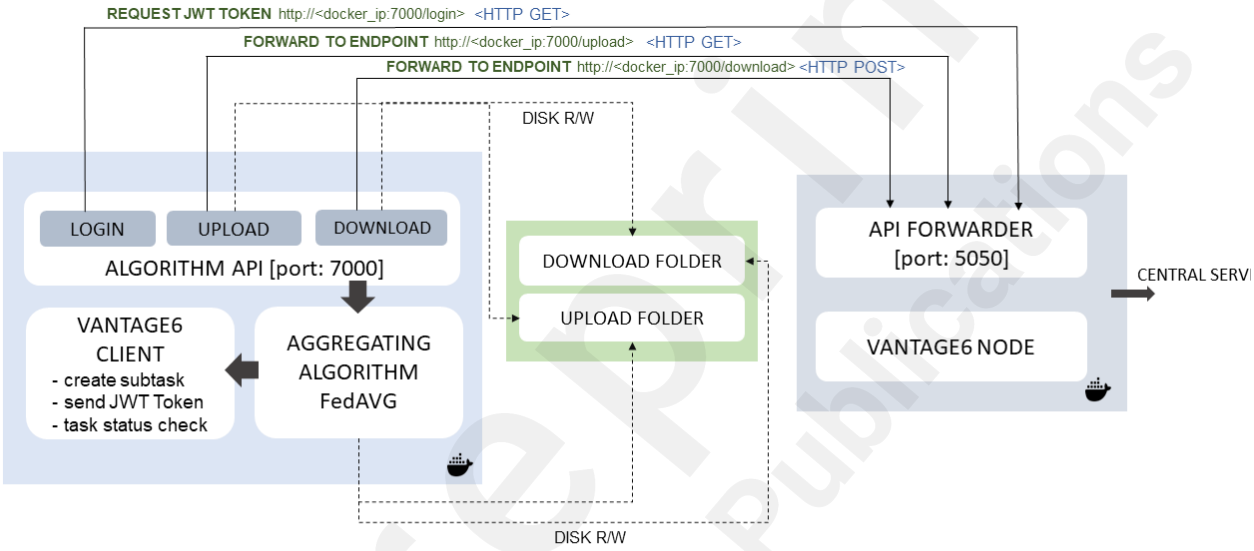

Figure 3: Architecture of the Secure Aggregation Server showing incoming and outgoing requests from the data station nodes. The upload and download folder are temporary locations used within the running docker container to store the local and averaged models through disk read/write operations. The API FORWARDER running at port 5050 and embedded within the Vantage6 infrastructure forwards the incoming requests from the data station nodes to the Algorithm API running at local port 7000 within the docker container through HTTP requests. The SAS is hosted behind the firewall of a proxy server which allows only Hypertext Transfer Protocol Secure (HTTPS) communication from the participating nodes.

Data Stations

Data stations are devices located within the confines of each hospital's jurisdiction that are not reachable or accessible from external sources other than Vantage6. The Data stations communicate with the central server through a pull mechanism. Furthermore, the data stations not only serve as hosts for the infrastructure node but also offer the essential computational resources required for training the deep learning network. The infrastructure node is the software component installed in the Data Stations that orchestrates the local execution of the model and its communication with the central server and the SAS. Each Data Stations is equipped with at least one Graphics Processing Unit (GPU), which enables the execution of Convolutional Neural Networks (CNNs). Pre-processing of the raw CT images was executed locally, using automated preprocessing scripts packaged as Docker containers and the preprocessed CT images are stored within a filesystem volume in each station. The CNN Docker is designed and allowed to access the preprocessed images during training. The

primary function of the data station is to receive instructions from both the SAS and the central server, perform the computations needed for training of the CNN algorithm, and subsequently transmit the model weights back to the respective sources. Figure 4 depicts the architectural layout of the data station and node component of the infrastructure.

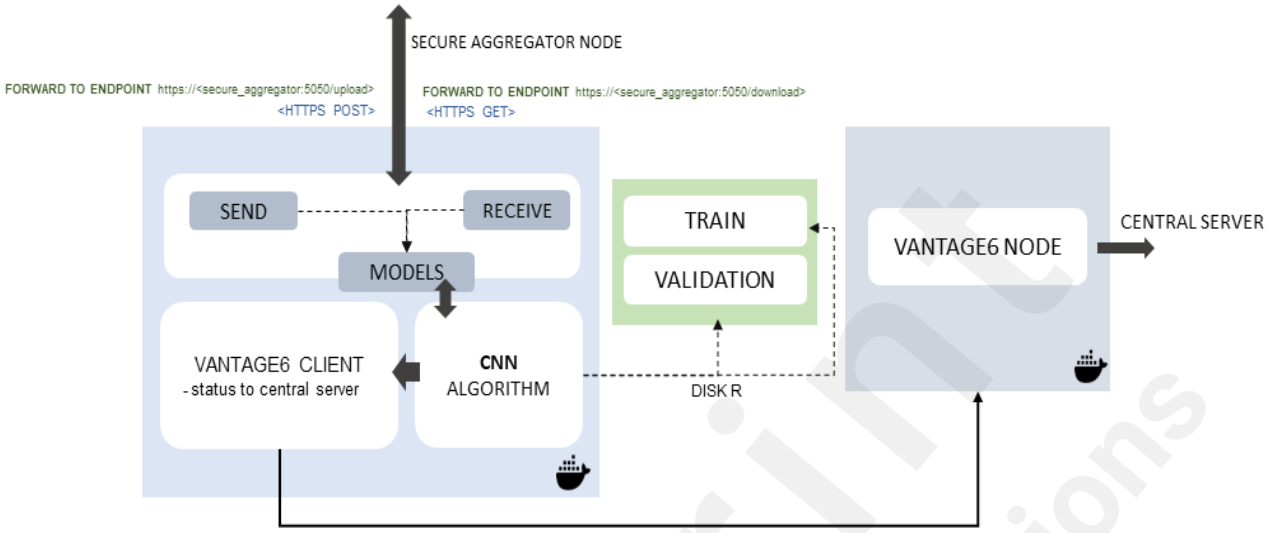

Figure 4: Architecture of the Data Station node component. The node runs the CNN algorithm to learn from the local data. The node further sends and receives model weights from the SAS. The TRAIN and VALIDATION folders are persistent locations withing the data stations storing the preprocessed NIFTI images. At the end of each training cycle the intermediate averaged model is first evaluated on the validation sample.

Train

The ‘Train’ in the form of a Docker image encompasses several components bundled together: an untrained U-Net[43,44], a type CNN architecture designed for image segmentation tasks for training on local data; the aggregation algorithm utilized for consolidating the models; and a secondary Python Flask API known as the *Algorithm API* for facilitating the communication of these models. The *Algorithm API* is designed to cater requests from the *API Forwarder* and is built within the algorithm container. Two levels of API ensured that the node could handle multiple requests and divert to appropriate docker containers. Furthermore, the first level of API also helps in restricting malicious requests by checking the JWT token signature, so that the models withing the master docker container are protected. Each data station is responsible for training and transmitting the CNN model to the aggregator server. This suggests that the aggregation algorithm exhibits a waiting period during which it ensures that all data stations have effectively transmitted their models to the server before proceeding for the next iterations The process is executed in an iterative manner until convergence is achieved or the the specified number of iterations is attained.

Tracks and Track Provider

The various infrastructure components establish coordination among themselves through the utilization of secure communication channels commonly referred to as ‘Tracks’. The communication channels are enabled with end-to-end encryption. The responsibility for the maintenance of the infrastructure, including the hosting of the central coordinating server and the specialized SAS,

lies with the track provider. The track provider is additionally accountable for the maintenance of the 'Tracks' and aids the Data Providers in establishing the local segment of the infrastructure known as 'Nodes'.

Data Provider

Data providers refer to hospitals and healthcare organizations that are responsible for curating the pertinent datasets utilized for training the deep learning network. The responsibility of hosting the Data Stations within their respective local jurisdiction lies with the Data Provider. They exercise authority over the data as well as the infrastructure component called the node.

Researcher

The researcher is responsible for activating the deep learning algorithm and engaging in the authentication process with the central coordinating server using a registered username and password. This allows the researcher to establish their identity and gain secure access to the system, with their communication safeguarded through end-to-end encryption. The researcher can then assign tasks to individual nodes, monitor progress, and terminate tasks in the event of failure. Importantly, the researcher's methodology is designed to keep the intermediate outcomes of the iterative deep learning training process inaccessible, ensuring that the ultimate global model can only be obtained upon completion of all training iterations, thereby mitigating the risk of unauthorized access by malicious researchers to the intermediate models and providing a security mechanism against insider attacks.

Training Process

Each of the components described above work in a coordinated manner to accomplish the convergence of the deep learning algorithm. The training process begins with the researcher authenticating with the central server. Upon successful authentication, the researcher specifies the task details, including a pre-built Docker image, input parameters, number of iterations, and the identity of the SAS (Secure Aggregation Server). The task is then submitted to the central server, which forwards it to the connected nodes. The SAS is the first to receive the task request. It downloads the specified Docker image from the registry and initiates the master algorithm. The master algorithm orchestrates the training at each data station node through the central server. The central server then forwards a sub-task request to all the data stations. Like the SAS, the data nodes download the same Docker image and initiate the node part of the algorithm. The node algorithm runs the learning process on local data for the specified number of epochs. After each training cycle, the node algorithm sends the local model weights to the SAS.

The SAS verifies the JWT signature of each received model and forwards the request to the *Algorithm API*. The *Algorithm API* extracts the weight and metadata information of the models. Once the SAS receives all the required locally trained models for that cycle, it initiates the FedAvg algorithm to consolidate the models and create an intermediate averaged model, which is stored locally. This completes the first iteration of the training cycle. For the second and subsequent iterations, the data stations request the SAS to send the intermediate averaged model weights from the previous iteration. The SAS

validates these requests and sends the model weights to the data stations, which then use them for further training on their local data. This cycle of training and averaging continues until the model converges or the desired number of iterations are reached.

At the end of the training process, the SAS sends a notification to the researcher indicating the successful completion of the task. The researcher can then download the final global model from the server. It is important to note that during the training iterations, the researcher or other users of the infrastructure do not have access to the intermediate averaged models generated by the SAS. This design choice prevents the possibility of insider attacks and data leakage, as users cannot regenerate patterns from the training data using the intermediate models. Figure 5 shows the diagrammatic representation of the training process spread across the infrastructure components.

Code Availability

The federated deep learning infrastructure and algorithm employed in this research are open-source and publicly available. The codebase, encompassing the components of the infrastructure, the algorithm, and wrappers for running it in the infrastructure and the researcher notebooks, are all available and deposited on GitHub, a public repository platform, under the Apache 2.0 license. This open access allows the research community to scrutinize and leverage our implementation for further development in the field of federated learning.

The Vantage6[45,27] (v2.0.0) open-source software was customized to cater to the specific requirements for running the deep learning algorithm. The central server (Vantage6 v2.0.0) and the aggregator server were hosted by Medical Data Works BV in two separate cloud virtual machines (Microsoft Azure). At each participating center the “node” component of the software was installed and setup either on a physical or virtual machine running Ubuntu v16.0 or above with an installation of Python, (v3.7 or above), Docker Desktop (personal edition) and CUDA GPU interface (v11.0). The source code of the customized “node”[46] and setup instructions[47] are available on respective GitHub repositories. The federated deep learning algorithm was adapted to the infrastructure as Python scripts[48] and wrapped in a docker container. Separately, the “researcher” notebooks[49] containing python scripts for connecting to the infrastructure and running the algorithms are also available on GitHub.

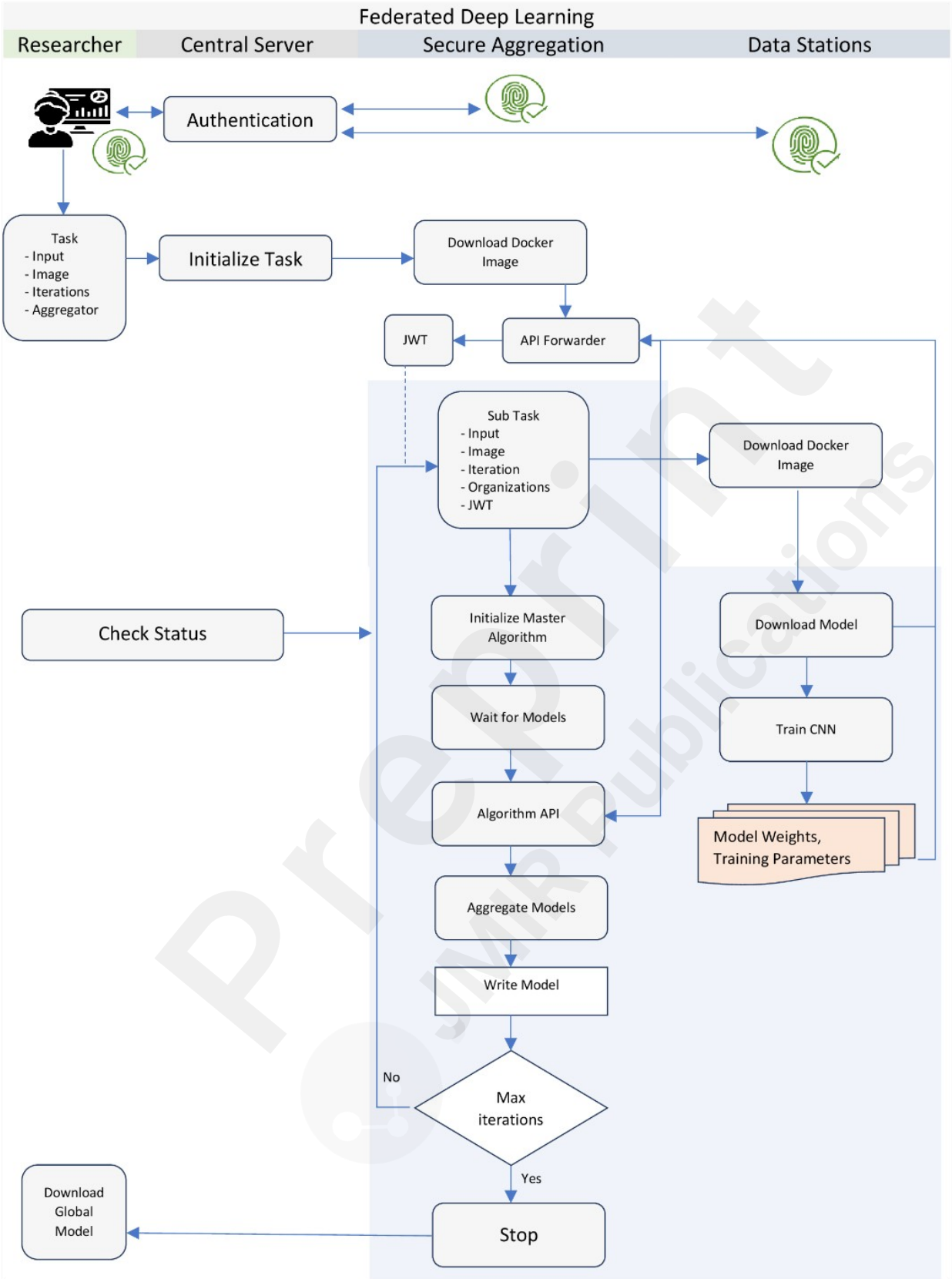

Figure 5: Process Illustration of Federated Deep Learning Training. All entities including the researcher, the central aggregation server, and the data stations at first authenticate with the central server. The researcher creates a task description and submits the task to the central server which then forwards the request to the secure aggregation node to start the master task. The master task then sends a request to all data stations to download the algorithm docker image and start training on the local data. At the end of each local training, the data stations send the models to the "API Forwarder" of the secure aggregation node by authenticating against a valid JWT Token. This request is forwarded to the secondary "Algorithm API", which writes the models to the disk from where the aggregation algorithm consolidates.

Results

The study was carried out and concluded in four primary stages using an agile approach – planning, design and development, partner recruitment, and execution of federated deep learning. The planning phase of the study, which encompassed a meticulous evaluation and determination of the following inquiries, held equal significance to the description of the clinical issue and data requirements.

- What are the minimum resource requirements for each participating center?
- How to design a safe and robust infrastructure to effectively address the requirements of a federated deep learning study?
- How can a reliable and data-agnostic federated deep learning algorithm be designed?
- What are the operational and logistical challenges associated with conducting a large-scale federated deep learning study?

The second phase i.e. the design and development phase primarily focused on the creation, testing, and customization of the Vantage6 infrastructure for studies specifically focused on deep learning. To meet the security demands of these investigations, the present study involved the development of the secure aggregation server, which was not originally included in the Vantage6 architecture. The CNN algorithm was packaged as a Docker container and made compatible with the Vantage6 infrastructure, allowing it to be easily deployed and utilized within the Vantage6 ecosystem. Prior to the deployment of the algorithm, it underwent testing using multiple test configurations consisting of data stations that were populated with public datasets.

The primary objective of the third phase entailed the recruitment of partners who displayed both interest and suitability from various global locations. The project consortium members became part of the project by obtaining the necessary Institutional Review Board (IRB) approvals and signing an infrastructure user agreement. This agreement enabled them to install the required infrastructure locally and carry out algorithmic execution. The inclusion criteria for patient data, as well as the technology used for data anonymization and preprocessing, were provided to each center. The team collaborated with each partner center to successfully implement the local component of the infrastructure.

The concluding stage of the study involved the simultaneous establishment of connections between all partner centers and the existing infrastructure. The algorithm was subsequently initiated by the researcher and the completion of the pre-determined set of federated iterations was awaited across all centers.

Proof of concept

The architectural strategy described above was implemented among ARGOS

consortium partners on real world lung cancer CT scans. For an initial “run-up” of the system, we deployed the abovementioned PHT system across 12 institutions, located in 8 countries and 4 continents. A list of members participating in the ARGOS consortium can be found on the online study protocol[26]. In total, 2078 subjects were accessible via the infrastructure for training (n=1606) and holdout validation (n=472). For this initial training experiment, the 12 centers were divided into two groups. The first, referred to as Group A, comprised seven collaborators, and we were able to reach a total of 64 iterations of model training each with 10000 steps per iteration. Likewise, Group B comprising 6 hospitals was able to train the deep learning model for 26 iterations. It was observed that no significant improvement of the model was observed for both groups after 26th iteration. The results from the proof-of-concept study are shown in the Figure 6 below.

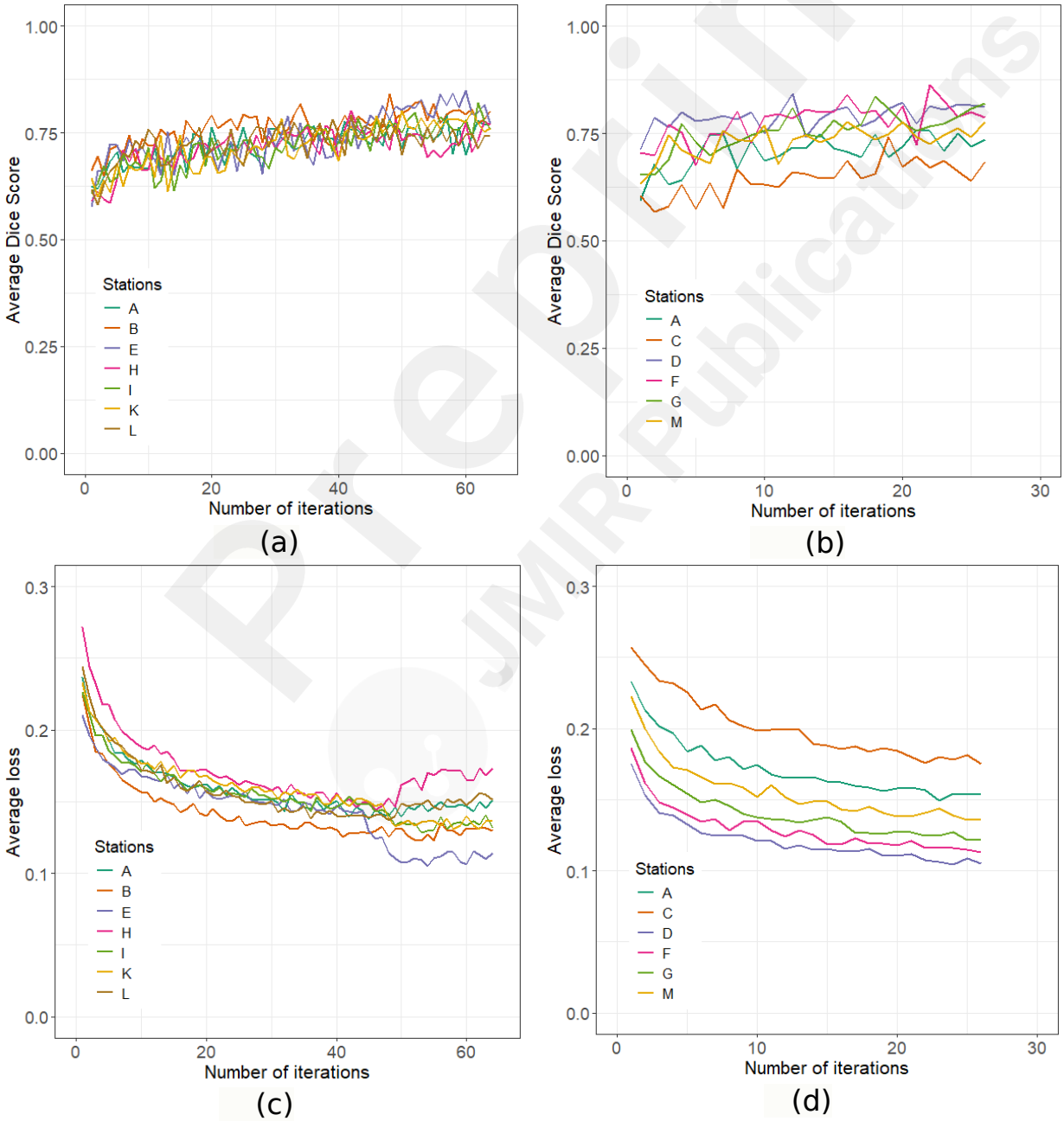

Figure 6: Plots showing the results from training the CNN on two groups - Group 1 (A, B, E, H, I, K, L) and Group 2 (A, C, D, F, G,

M). (a) Average dice score per iteration of the model trained on Group 1. (b) Average dice score per iteration of the model trained on Group 2. (c) Average training loss per iteration of the model trained on Group 1 (d) Average training loss of the model trained on Group 2.

While the training time for the models was similar at each center, how quickly they could be uploaded and downloaded depended heavily on the quality of the internet connection. This meant the entire process was significantly slowed down by the center with the slowest internet. While the training time for the models was similar at each center, how quickly they could be uploaded and downloaded depended heavily on the quality of the internet connection. This meant the entire process was significantly slowed down by the center with the slowest internet.

Conclusion and Discussion

One of the most used methodologies in recent years has been the use of Federated Learning for promoting research on privacy sensitive data. To orchestrate federated learning on non-structured data in the horizontal partitioning context, it is essential to develop specialized software for edge computation and technical infrastructures for cloud aggregation. These infrastructures enable FML responsibilities to be carried out in a secure and regulated manner. However, only a limited number of these studies have documented the background governance strategies and the ELSI framework for conducting such studies.

The study presented a novel approach for executing large-scale federated deep learning on medical imaging data, integrating geographically dispersed real-world patient data from cross-continental hospital sites. The deep learning algorithm was designed to automatically delineate the gross tumor volume from chest CT images of lung cancer patients who underwent radiotherapy treatment. The underlying federated learning infrastructure architecture was designed to securely perform deep learning training and was tested for vulnerabilities from known security threats. This paper predominantly discussed the federated learning infrastructure architecture and presented a firsthand experience of conducting such studies. The preliminary training of deep learning algorithm serves as the feasibility demonstration of the methodology, and further refinement is required to achieve an acceptable clinical-grade accuracy and generalizability.

The study employed an open source and freely accessible technological stack to demonstrate the feasibility and applicability of federated deep learning. Vantage6, a Python-based federated learning infrastructure, is used to train and co-ordinate deep learning execution. TensorFlow and Flask, both open-source Python libraries, are employed for the development of the algorithm, subsequently encapsulated within Docker services for containerization purposes. The communication channels between the hospital, central server and the aggregation node have been secured using HTTPS and Secure Hash Algorithm (SHA) encryption. The hospital sites' computer systems were based on the Ubuntu operating system and equipped with at least one GPU to enhance computational capabilities. The participating centers had the flexibility to choose any CUDA compatible GPU devices and determine the number of GPUs

to use, enabling resource-constrained centers to contribute. However, a limitation exists in terms of computational time due to the synchronous training process being dependent on the slowest participant.

The infrastructure has been tested against known security attacks and as defined by the OWASP top-ten categories[50]. It has been found that the Vantage6 application is impeccable against Insecure Design, Software and Data Integrity Failures, Security Logging and Monitoring Failures, Server-Side Request Forgery and sufficiently secured against Broken Access Control, Cryptographic Failures, Injection, Security Misconfigurations, Vulnerable and Outdated Components and finally Identification and Authentication Failures. Since the infrastructure is dependent on other underlying technology like Docker and Flask-API, the security measures in these technologies also affect the overall security of the infrastructure. Additionally, the infrastructure is hosted behind proxy firewalls, adding to its overall security against external threats.

In this study we implemented a SAS positioned between the data nodes (e.g., hospitals, clinics) and the central server. The SAS plays a crucial role in strengthening the privacy and confidentiality of the learning process. The SAS acts as an intermediary that temporarily stores the local model updates from the participating data nodes, ensuring complete isolation from the central server, researchers, and any external intruders. The key benefits of using a dedicated SAS over a random aggregation mechanism in FL are:

- Privacy protection of individual user data and model updates:
    - The secure aggregation protocol ensures that the central server only learns the aggregated sum of all user updates, without being able to access or infer the individual user's private data or model updates.
    - By isolating the intermediate updates, the secure aggregation process prevents external attackers from performing model inversion attacks.
- Tolerance to user dropouts:
    - The secure aggregation server is designed to handle situations where some users fail to complete the execution. In the case of synchronous training, the server stores the latest successful model, enabling data nodes to pick up where they left off instead of restarting from scratch.
- Integrity of the aggregation process:
    - The secure aggregation protocol provides mechanisms to verify the integrity of the intermediate models by allowing only the known data nodes to send a model. This maintains the reliability and trustworthiness of the federated learning system.

Federated learning offers two main approaches for model aggregation: sending

gradients or weights[51,52]. In gradient sharing, data nodes update local models and transmit the gradients of their parameters for aggregation. Conversely, weight sharing involves sending the fully updated model weights directly to the server for aggregation. Sharing gradients has a higher risk of model inversion attacks. In the study presented here, the data nodes sent model weights instead of model gradients, thus preventing "gradient leakage" problem. However, weight sharing isn't failproof either[53]. And the secure aggregation server plays a crucial role again in preventing users – internal or external from accessing the weights from the aggregator machine.

The deployment of the FL infrastructure and training of the deep learning algorithm presented unique challenges that needed to be catered. Some of them are listed below –

- Heterogeneity across hospitals: Initially, it was not possible to confirm the technology environment at each site. This required significant work to overcome the obstacles connected with each center while deploying a functional infrastructure, good communication, and efficient algorithms.

- Inconsistent IT Policies: Standardizing the setup across institutions was hindered by varying IT governance and network regulations in different healthcare systems across different countries.

- Clinical Expertise Gap: The predominance of medical personnel over IT specialists at participating hospitals necessitated extensive documentation to ensure clinician comprehension of the federated learning process.

- Network Bottlenecks: Network configurations at participating sites significantly impacted training duration, often leading to delays in model convergence.

The study presented in the paper has identified several areas that require further investigation and improvements. While the findings are valuable, the infrastructure, algorithm and the processes still need to be made more secure, private, trustworthy, robust, and seamless. For example, incorporating homomorphic encryption of the learned models will enhance privacy and provide model obfuscation against inversion attacks. Finally, to further enhance confidence and trust in federated artificial intelligence, it is crucial to conduct additional studies involving a larger number of participating centers and a thorough clinical evaluation of the models.

Funding

The authors Ananya Choudhury, Leroy Volmer, Rianne Fijten and Leonard Wee acknowledge financial support from the Dutch Research Council (NWO) (TRAIN project, dossier 629.002.212) and the Hanarth Foundation.

Acknowledgement

We would like to express our sincere appreciation and gratitude to Integraal

Kankercentrum Nederland (IKNL), Netherlands, for their invaluable contribution in providing us with the necessary infrastructure support. We express our gratitude to Medical Data Works, Netherlands for their role as the infrastructure service provider in hosting the central and secure aggregation server. We also express our gratitude to Varsha Gouthamchand and Sander Puts for their contribution in successful execution of the experiments. In conclusion, we express our gratitude to the various data-providing organizations for their substantial support and collaboration throughout all stages of the project.

Conflict of Interest

The authors Dr. Andre Dekker and Johan van Soest are both co-founders, shareholders, and directors of Medical Data Works B.V.

1. Sun C, Ippel L, Dekker A, Dumontier M, van Soest J. A systematic review on privacy-preserving distributed data mining. *Data Science*. 2021;4(2):121-150. doi:10.3233/DS-210036

2. Choudhury A, Sun C, Dekker A, Dumontier M, van Soest J. Privacy-Preserving Federated Data Analysis: Data Sharing, Protection, and Bioethics in Healthcare. In: El Naqa I, Murphy MJ, eds. *Machine and Deep Learning in Oncology, Medical Physics and Radiology*. Springer International Publishing; 2022:135-172. doi:10.1007/978-3-030-83047-2_8

3. Deist TM, Dankers FJWM, Ojha P, et al. Distributed learning on 20 000+ lung cancer patients – The Personal Health Train. *Radiotherapy and Oncology*. 2020;144:189-200. doi:10.1016/j.radonc.2019.11.019

4. Choudhury A, Theophanous S, Lønne PI, et al. Predicting outcomes in anal cancer patients using multi-centre data and distributed learning – A proof-of-concept study. *Radiotherapy and Oncology*. 2021;159:183-189. doi:10.1016/j.radonc.2021.03.013

5. Beyan O, Choudhury A, van Soest J, et al. Distributed Analytics on Sensitive Medical Data: The Personal Health Train. *Data Intelligence*. 2020;2(1-2):96-107. doi:10.1162/dint_a_00032

6. Moncada-Torres A, Martin F, Sieswerda M, Van Soest J, Geleijnse G. VANTAGE6: an open source priVAcy preserviNg federaTed leArninG infrastructurE for Secure Insight eXchange. *AMIA Annu Symp Proc*. 2020;2020:870-877.

7. Becker R, Chokoshvili D, Comandé G, et al. Secondary Use of Personal Health Data: When Is It “Further Processing” Under the GDPR, and What Are the Implications for Data Controllers? *European Journal of Health Law*. 2022;30(2):129-157. doi:10.1163/15718093-bja10094

8. El Naqa I, Ruan D, Valdes G, et al. Machine learning and modeling: Data, validation, communication challenges. *Medical Physics*. 2018;45(10):e834-e840. doi:10.1002/mp.12811

9. van Stiphout R, Deist TM, Walsh S, et al. How to Share Data and Promote a Rapid Learning Health Medicine? In: Valentini V, Schmoll HJ, van de Velde CJH, eds. *Multidisciplinary Management of Rectal Cancer: Questions and Answers*. Springer International Publishing; 2018:623-634. doi:10.1007/978-3-319-43217-5_74

10. Kazmierska J, Hope A, Spezi E, et al. From multisource data to clinical decision aids in

radiation oncology: The need for a clinical data science community. *Radiotherapy and Oncology*. 2020;153:43-54. doi:10.1016/j.radonc.2020.09.054

11. Fischer-Hübner S. Privacy-Enhancing Technologies. In: LIU L, ÖZSU MT, eds. *Encyclopedia of Database Systems*. Springer US; 2009:2142-2147. doi:10.1007/978-0-387-39940-9_271

12. Usage Patterns of Privacy-Enhancing Technologies | Proceedings of the 2020 ACM SIGSAC Conference on Computer and Communications Security. Accessed April 25, 2024. https://dl.acm.org/doi/10.1145/3372297.3423347

13. Emerging privacy-enhancing technologies: Current regulatory and policy approaches | en | OECD. Accessed April 25, 2024. https://www.oecd.org/publications/emerging-privacy-enhancing-technologies-bf121be4-en.htm

14. Kavianpour S, Sutherland J, Mansouri-Benssassi E, Coull N, Jefferson E. Next-Generation Capabilities in Trusted Research Environments: Interview Study. *J Med Internet Res*. 2022;24(9):e33720. doi:10.2196/33720

15. PageWriter-MSFT. Secure research environment for regulated data - Azure Architecture Center. Accessed April 25, 2024. https://learn.microsoft.com/en-us/azure/architecture/ai-ml/architecture/secure-compute-for-research

16. Imaging Data Commons | CRDC. Accessed April 25, 2024. https://datacommons.cancer.gov/repository/imaging-data-commons

17. Kotter E, Marti-Bonmati L, Brady AP, Desouza NM, European Society of Radiology (ESR). ESR white paper: blockchain and medical imaging. *Insights into Imaging*. 2021;12(1):82. doi:10.1186/s13244-021-01029-y

18. Sultana M, Hossain A, Laila F, Taher KA, Islam MN. Towards developing a secure medical image sharing system based on zero trust principles and blockchain technology. *BMC Medical Informatics and Decision Making*. 2020;20(1):256. doi:10.1186/s12911-020-01275-y

19. PHT_Manifesto.pdf. Accessed May 2, 2024. https://www.dtls.nl/wp-content/uploads/2017/12/PHT_Manifesto.pdf

20. McMahan B, Moore E, Ramage D, Hampson S, Arcas BA y. Communication-Efficient Learning of Deep Networks from Decentralized Data. In: *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics*. PMLR; 2017:1273-1282. Accessed June 29, 2023. https://proceedings.mlr.press/v54/mcmahan17a.html

21. Zhang C, Choudhury A, Shi Z, et al. Feasibility of Privacy-Preserving Federated Deep Learning on Medical Images. *International Journal of Radiation Oncology, Biology, Physics*. 2020;108(3):e778. doi:10.1016/j.ijrobp.2020.07.234

22. Choudhury A, van Soest J, Nayak S, Dekker A. Personal Health Train on FHIR: A Privacy Preserving Federated Approach for Analyzing FAIR Data in Healthcare. In: Bhattacharjee A, Borgohain SKr, Soni B, Verma G, Gao XZ, eds. *Machine Learning, Image Processing, Network Security and Data Sciences*. Communications in Computer and Information Science. Springer;

2020:85-95. doi:10.1007/978-981-15-6315-7_7

23. Gouthamchand V, Choudhury A, Hoebers FJP, et al. Making head and neck cancer clinical data Findable-Accessible-Interoperable-Reusable to support multi-institutional collaboration and federated learning. *BJR|Artificial Intelligence*. 2024;1(1):ubae005. doi:10.1093/bjrai/ubae005

24. Sun C, van Soest J, Koster A, et al. Studying the association of diabetes and healthcare cost on distributed data from the Maastricht Study and Statistics Netherlands using a privacy-preserving federated learning infrastructure. *J Biomed Inform*. 2022;134:104194. doi:10.1016/j.jbi.2022.104194

25. Medical Data Works. Accessed April 25, 2024. https://www.medicaldataworks.nl/governance

26. Dekker A. *ARtificial Intelligence for Gross Tumour vOlume Segmentation*. clinicaltrials.gov; 2024. Accessed January 1, 2024. https://clinicaltrials.gov/study/NCT05775068

27. vantage6 documentation. Accessed April 25, 2024. https://docs.vantage6.ai/en/main/

28. Tao Z, Lyu S. A Survey on Automatic Delineation of Radiotherapy Target Volume based on Machine Learning. *Data Intelligence*. Published online February 11, 2023:1-22. doi:10.1162/dint_a_00204

29. Liu X, Li KW, Yang R, Geng LS. Review of Deep Learning Based Automatic Segmentation for Lung Cancer Radiotherapy. *Front Oncol*. 2021;11:717039. doi:10.3389/fonc.2021.717039

30. Ma Y, Mao J, Liu X, et al. Deep learning-based internal gross target volume definition in 4D CT images of lung cancer patients. *Med Phys*. 2023;50(4):2303-2316. doi:10.1002/mp.16106

31. Zhang F, Wang Q, Li H. Automatic Segmentation of the Gross Target Volume in Non-Small Cell Lung Cancer Using a Modified Version of ResNet. *Technol Cancer Res Treat*. 2020;19:1533033820947484. doi:10.1177/1533033820947484

32. Xie H, Chen Z, Deng J, Zhang J, Duan H, Li Q. Automatic segmentation of the gross target volume in radiotherapy for lung cancer using transresSEUnet 2.5D Network. *Journal of Translational Medicine*. 2022;20(1):524. doi:10.1186/s12967-022-03732-w

33. Riedel P, von Schwerin R, Schaudt D, Hafner A, Späte C. ResNetFed: Federated Deep Learning Architecture for Privacy-Preserving Pneumonia Detection from COVID-19 Chest Radiographs. *J Healthc Inform Res*. 2023;7(2):1-22. doi:10.1007/s41666-023-00132-7

34. Nazir S, Kaleem M. Federated Learning for Medical Image Analysis with Deep Neural Networks. *Diagnostics (Basel)*. 2023;13(9):1532. doi:10.3390/diagnostics13091532

35. Shiri I, Vafaei Sadr A, Akhavan A, et al. Decentralized collaborative multi-institutional PET attenuation and scatter correction using federated deep learning. *Eur J Nucl Med Mol Imaging*. 2023;50(4):1034-1050. doi:10.1007/s00259-022-06053-8

36. Zhang M, Qu L, Singh P, Kalpathy-Cramer J, Rubin DL. SplitAVG: A Heterogeneity-Aware Federated Deep Learning Method for Medical Imaging. *IEEE J Biomed Health Inform*. 2022;26(9):4635-4644. doi:10.1109/JBHI.2022.3185956

37. Shiri I, Vafaei Sadr A, Amini M, et al. Decentralized Distributed Multi-institutional PET

Image Segmentation Using a Federated Deep Learning Framework. *Clin Nucl Med*. 2022;47(7):606-617. doi:10.1097/RLU.0000000000004194

38. Sarma KV, Harmon S, Sanford T, et al. Federated learning improves site performance in multicenter deep learning without data sharing. *J Am Med Inform Assoc*. 2021;28(6):1259-1264. doi:10.1093/jamia/ocaa341

39. Harmon SA, Sanford TH, Xu S, et al. Artificial intelligence for the detection of COVID-19 pneumonia on chest CT using multinational datasets. *Nat Commun*. 2020;11(1):4080. doi:10.1038/s41467-020-17971-2

40. Durga R, Poovammal E. FLED-Block: Federated Learning Ensembled Deep Learning Blockchain Model for COVID-19 Prediction. *Front Public Health*. 2022;10:892499. doi:10.3389/fpubh.2022.892499

41. Pati S, Baid U, Edwards B, et al. Federated learning enables big data for rare cancer boundary detection. *Nat Commun*. 2022;13(1):7346. doi:10.1038/s41467-022-33407-5

42. Patil RB. *Prognostic and Prediction Modelling with Radiomics for Non-Small Cell Lung Cancer*. Maastricht University; 2020. doi:10.26481/dis.20201006rp

43. Oreiller V, Andrearczyk V, Jreige M, et al. Head and neck tumor segmentation in PET/CT: The HECKTOR challenge. *Med Image Anal*. 2022;77:102336. doi:10.1016/j.media.2021.102336

44. Iantsen A, Jaouen V, Visvikis D, Hatt M. Squeeze-and-Excitation Normalization for Brain Tumor Segmentation. In: Crimi A, Bakas S, eds. *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*. Springer International Publishing; 2021:366-373. doi:10.1007/978-3-030-72087-2_32

45. IKNL/vantage6: Docker CLI package for the vantage6 infrastructure. Accessed May 1, 2024. https://github.com/IKNL/vantage6/tree/DEV3

46. Martin F, Beusekom B van, Sieswerda M, et al. IKNL/vantage6: vantage6 ARGOS. Published online May 6, 2024. doi:10.5281/zenodo.11121679

47. MaastrichtU-CDS/argos-infrastructure. Published online May 12, 2023. Accessed May 1, 2024. https://github.com/MaastrichtU-CDS/argos-infrastructure

48. MaastrichtU-CDS/projects_argos_argos-code-repo_full-algorithm. GitHub. Accessed May 1, 2024. https://github.com/MaastrichtU-CDS/projects_argos_argos-code-repo_full-algorithm

49. MaastrichtU-CDS/projects_argos_argos-code-repo_researcher-notebooks. GitHub. Accessed May 1, 2024. https://github.com/MaastrichtU-CDS/projects_argos_argos-code-repo_researcher-notebooks

50. OWASP Top Ten | OWASP Foundation. Accessed May 2, 2024. https://owasp.org/www-project-top-ten/

51. Moshawrab M, Adda M, Bouzouane A, Ibrahim H, Raad A. Reviewing Federated Learning Aggregation Algorithms; Strategies, Contributions, Limitations and Future Perspectives. *Electronics*. 2023;12(10):2287. doi:10.3390/electronics12102287

52. Liu P, Xu X, Wang W. Threats, attacks and defenses to federated learning: issues, taxonomy and perspectives. *Cybersecurity*. 2022;5(1):4. doi:10.1186/s42400-021-00105-6

53. Boenisch F, Dziedzic A, Schuster R, Shamsabadi AS, Shumailov I, Papernot N. When the Curious Abandon Honesty: Federated Learning Is Not Private. In: *2023 IEEE 8th European Symposium on Security and Privacy (EuroS&P)*. ; 2023:175-199. doi:10.1109/EuroSP57164.2023.00020

List of Abbreviations

GTV	Gross Tumor Volume
FL	Federated Learning
AI	Artificial Intelligence
PHT	Personal Health Train
CT	Computed Tomography
GDPR	General Data Protection Regulations
HIPAA	Health Insurance Portability and Accountability Act of 1996
PET	Privacy Enhancing Technologies
SRE	Secure Research Environment
NIH	National Institutes of Health
ARGOS	ARtificial intelligence for Gross tumor vOlume Segmentation
ELSI	Ethical Legal Societal Implications
DICOM	Digital Imaging and Communications in Medicine
SAS	Secure Aggregation Server
REST	Representational State Transfer
API	Application Programming Interface
JWT	JSON Web Token
GPU	Graphical Processing Unit
CNN	Convolutional Neural Network
HTTPS	HyperText Transfer Protocol Secure
SHA	Secure Hash Algorithm