Article

Patterns

A federated learning architecture for secure and private neuroimaging analysis

Highlights

- Federated learning can perform comparably with respect to centralized neuroimaging analysis
- Training with fully homomorphic encryption (FHE) in MetisFL has low overhead (∼7%)
- MetisFL is secure: FHE protects against outsiders and gradient noise against insiders
- MetisFL controller optimizations lead to 10-fold reduction of federated training time

Authors

Dimitris Stripelis, Umang Gupta,
Hamza Saleem, ...,
Muhammad Naveed,
Paul M. Thompson, José Luis Ambite

Correspondence

stripeli@isi.edu

In brief

This research presents MetisFL, a scalable, secure, and private federated learning architecture, highlighting its importance in neuroimaging analysis by comparing deep learning models trained in centralized and federated settings, including Alzheimer’s disease prediction and brain age gap estimation from MRI scans. To protect the federation against outsider attacks, MetisFL uses fully homomorphic encryption, optimized to have low (∼7% runtime) overhead. To protect against insider model inversion and membership inference attacks, MetisFL trains under information-theoretic gradient noise.

Stripelis et al., 2024, Patterns *5*, 101031
August 9, 2024 © 2024 The Author(s). Published by Elsevier Inc.
https://doi.org/10.1016/j.patter.2024.101031

Patterns

CellPress
OPEN ACCESS

Article

A federated learning architecture for secure and private neuroimaging analysis

Dimitris Stripelis,[1,2,5,6,*] Umang Gupta,[1,2,5] Hamza Saleem,[2,5] Nikhil Dhinagar,[3] Tanmay Ghai,[1,2]
Chrysovalantis Anastasiou,[2] Rafael Sánchez,[1,2] Greg Ver Steeg,[4] Srivatsan Ravi,[1,2] Muhammad Naveed,[2]
Paul M. Thompson,[3] and José Luis Ambite[1,2]
[1]University of Southern California, Information Sciences Institute, Marina del Rey, CA 90292, USA
[2]University of Southern California, Computer Science Department, Los Angeles, CA 90089, USA
[3]University of Southern California, Imaging Genetics Center, Stevens Neuroimaging and Informatics Institute, Keck School of Medicine, Marina del Rey, CA 90292, USA
[4]University of California, Riverside, Riverside, CA 92521, USA
[5]These authors contributed equally
[6]Lead contact
*Correspondence: stripeli@isi.edu

https://doi.org/10.1016/j.patter.2024.101031

THE BIGGER PICTURE Federated learning (FL) has emerged as a promising machine learning approach to enable large-scale analyses across multiple institutions without sharing data. With FL, data remain private and secure. Data are never shared; only encrypted model parameters are shared and aggregated. This research exemplifies the potential of FL consortia, critically comparing models trained centrally and in a federation in the neuroimaging domain. We show that FL performs comparably with respect to centralized approaches and outperforms them in cases where more sites join the federation than are willing to share data. Ultimately, we want to raise awareness of the advantages and challenges of FL among biomedical and healthcare researchers and encourage scientists to establish ever-larger federated consortia to improve biomedical analysis in real-world settings.

SUMMARY

The amount of biomedical data continues to grow rapidly. However, collecting data from multiple sites for joint analysis remains challenging due to security, privacy, and regulatory concerns. To overcome this challenge, we use federated learning, which enables distributed training of neural network models over multiple data sources without sharing data. Each site trains the neural network over its private data for some time and then shares the neural network parameters (i.e., weights and/or gradients) with a federation controller, which in turn aggregates the local models and sends the resulting community model back to each site, and the process repeats. Our federated learning architecture, MetisFL, provides strong security and privacy. First, sample data never leave a site. Second, neural network parameters are encrypted before transmission and the global neural model is computed under fully homomorphic encryption. Finally, we use information-theoretic methods to limit information leakage from the neural model to prevent a "curious" site from performing model inversion or membership attacks. We present a thorough evaluation of the performance of secure, private federated learning in neuroimaging tasks, including for predicting Alzheimer's disease and for brain age gap estimation (BrainAGE) from magnetic resonance imaging (MRI) studies in challenging, heterogeneous federated environments where sites have different amounts of data and statistical distributions.

INTRODUCTION

Deep learning and traditional machine learning methods are now widely applied across biomedical research.[1] These methods have been particularly successful in medical imaging,[2] including image reconstruction and enhancement,[3] automated segmentation and labeling of key structures,[4] computer-aided diagnosis,[5] pathology detection,[6] disease subtyping,[7,8] and predictive analytics (e.g., modeling future recovery or decline).[9]

In neuroimaging, there has been great progress in automated diagnostic classification and subtyping of diseases, such as in Alzheimer's disease (AD) and Parkinson's disease, to assist in patient management and monitoring and to screen patients for eligibility for clinical trials. Some recent magnetic resonance imaging (MRI)-based classifiers have merged data from over 80,000 individuals for diagnostic classification.[10] The performance of deep learning methods depends heavily on the availability of large amounts of training data. Unfortunately, data acquisition is expensive in many areas of biomedical research, such as neuroimaging. Therefore, any organization or research group can collect only limited data.

To increase the amount of data to improve the statistical model's learning performance, research groups join together into consortia.[11] However, the need to protect patient data makes data sharing very challenging. Regulatory frameworks, such as the Health Insurance Portability and Accountability Act (HIPAA), require strict protection of health records and data collected for medical research. Privacy laws have spurred research into anonymization methods, for example, algorithms to remove facial information from MRI scans.[12–14] The inherent complexity and cost of enforcing security and privacy results in few large-scale data sharing efforts. Even when large consortia are established, they often perform only meta-analysis using traditional statistical methods instead of joint mega-analysis using deep learning methods. A paradigmatic example of large-scale meta-analysis is the ENIGMA Consortium.[11]

Federated learning[15–19] has emerged as a promising distributed approach for performing large-scale cross-institutional data analysis without moving the data out of their original location. Using federated learning, institutions can collaboratively apply classical statistical methods and train machine and deep learning models by aggregating the parameters (e.g., weights) of models trained on institutions' private data. Since subject data are never shared and parameters can be protected through encryption, privacy concerns are ameliorated. Federated learning is being increasingly applied in biomedical and healthcare domains.[19–23] Similarly, in the neuroimaging domain, many recent works have demonstrated the promise of federated learning for enabling federated neuroimaging data analysis in multisite consortia without data sharing.[24–28]

This paper demonstrates the potential of federated learning to accelerate and improve research outcomes through decentralized biomedical consortia. We conduct our analysis using MetisFL, our secure and private federated learning system. The MetisFL source code and the definitions of the models used to conduct the experiments in this study are publicly available on GitHub (https://github.com/bioint/MetisFL) and Zenodo[29] (https://dx.doi.org/10.5281/zenodo.11411754). Our design is modular, and extensible, and supports a variety of federated training policies. Here, we present a comprehensive description of the MetisFL framework and provide a qualitative comparison with other existing federated learning frameworks. A systematic quantitative evaluation appears in Stripelis et al.[30]

The MetisFL architecture[26] appears in Figure 1. Each site trains the neural network over its private data for some time and then shares the neural network parameters (i.e., weights and/or gradients) with a federation controller, which in turn aggregates the local models and sends the resulting community model back to each site, and the process repeats. Federated training in MetisFL is secure. Data are never shared. Model parameters are transmitted through secure communication channels. Moreover, model parameters are encrypted, and the global model is computed under fully homomorphic encryption (FHE) (using CKKS[31]), so even if the controller were compromised, the global model could not be attacked. Finally, we use information-theoretic methods to limit information leakage from the neural model to prevent a "curious" site within the federation from performing model inversion[32,33] or membership inference[34,35] attacks.

We present a thorough evaluation of the performance of secure, private federated learning in neuroimaging tasks, including predicting AD and making a brain age gap estimation (BrainAGE) from MRI studies, in challenging, heterogeneous federated environments where sites have different amounts of data and statistical distributions. Specifically, we show that research consortia based on federated learning, without data sharing, can achieve comparable learning performance with respect to centralized consortia, where data are aggregated into a single site. We show that our homomorphic encryption (HE) methods are practical, having little runtime overhead over unencrypted training. We show defense mechanisms against attacks on federated neural models and the trade-offs between security and learning performance. In summary, secure federated learning enables large, decentralized analysis of biomedical data without the burdens of data sharing. Since the performance of deep learning models improves with the amount of data used for training, federated learning over research consortia promises improvements in disease diagnosis, prognosis, biomarker detection, and many other advances in biomedical research.

RESULTS

Federated learning can perform comparably with respect to centralized analysis and even outperform it in the likely scenario where more sites are willing to join a federation than are willing to share data

We evaluated centralized and federated learning on two challenging neuroimaging learning tasks: the BrainAGE regression task and the AD detection classification task, both on structural MRI inputs.

For the BrainAGE task, we selected 10,446 MRI scans of healthy individuals (no neurological or psychiatric diagnoses) from the UK Biobank (UKBB).[36] Figure S1 shows the distribution of the UKBB samples. We trained and tested a 3D convolutional neural network (CNN) to predict BrainAGE on diverse centralized and federated environments. Figure S3 shows the 3D-CNN model architecture.

Figure 2 shows the performance of the centralized and federated models in terms of mean absolute error (MAE, y axis) and wall-clock time execution (x axis). Figure S4 shows model convergence based on communication cost. We tested four heterogeneous federated environments with different amounts of data per site (uniform, equal number of training samples per site, and skewed, a decreasing amount of training samples for each site) and different data distributions per site (IID, independent and identically distributed, where the local data distribution of each site is similar to the global distribution, and non-IID, where it differs). Figure 3 shows the different data distributions of the federated sites, and the insets in the plots in Figure 2 show the amounts of data per site. We report the results of

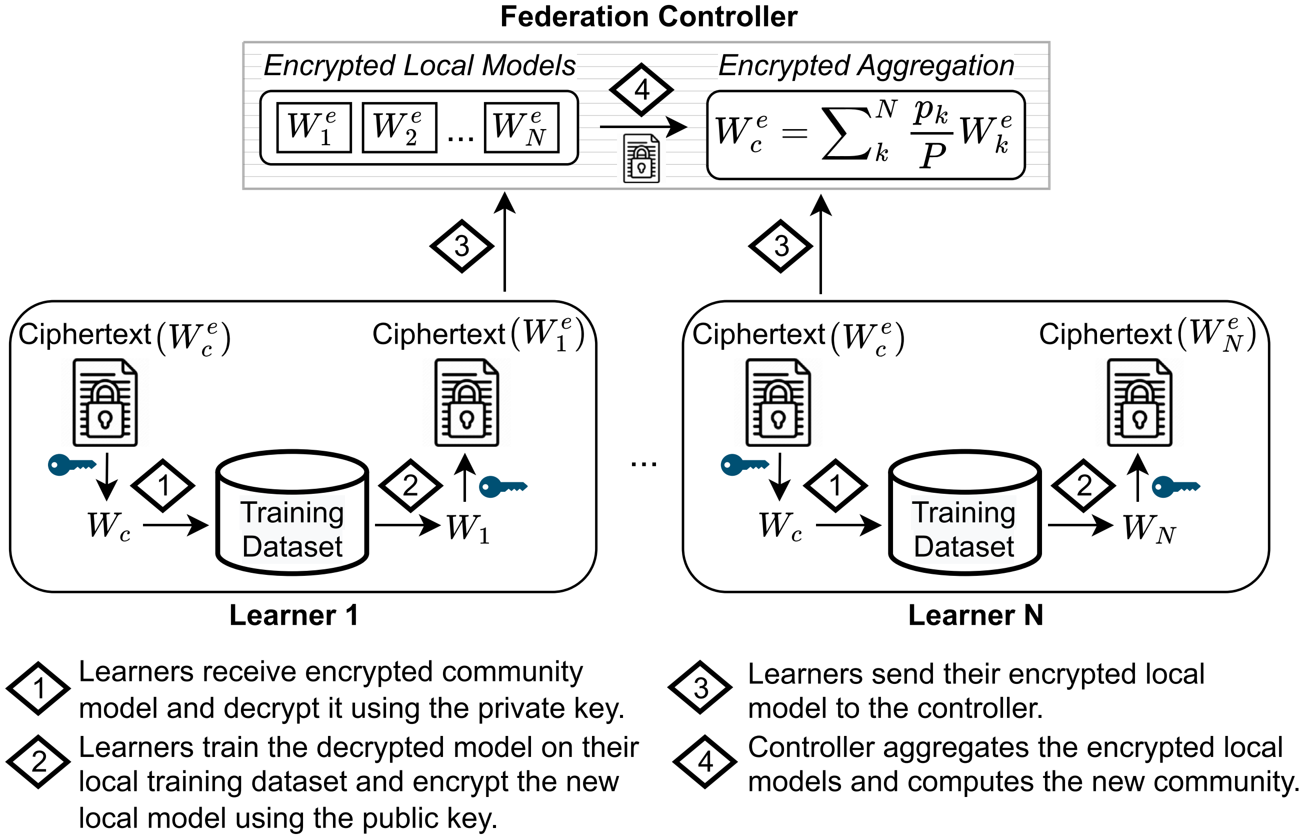

Federation Controller
Encrypted Local Models
Encrypted Aggregation
W_1^e
W_2^e
...
W_N^e
4
$W_c^e = \sum_k^N \frac{p_k}{P} W_k^e$
3
3
Ciphertext (W_c^e)
Ciphertext (W_1^e)
W_c
1
Training Dataset
2
W_1
Learner 1
...
Ciphertext (W_c^e)
Ciphertext (W_N^e)
W_c
1
Training Dataset
2
W_N
Learner N
1 Learners receive encrypted community model and decrypt it using the private key.
2 Learners train the decrypted model on their local training dataset and encrypt the new local model using the public key.
3 Learners send their encrypted local model to the controller.
4 Controller aggregates the encrypted local models and computes the new community.

Figure 1. MetisFL, a secure federated learning architecture
Steps 1–4 demonstrate the execution flow of federated training with encryption.

centralized training as constant horizontal lines to visualize the performance gap between centralized and federated models (centralized models converge much faster). Federated training achieves the same performance as centralized training in the uniform and IID environment. In the harder skewed or non-IID environments, there is a (small) gap between centralized and federated analysis when the same amount of data is available to both approaches (centralized [100%] in Figure 2).

However, the promise of federated learning is that federated consortia, which do not require data sharing, can enroll more sites than centralized consortia, which require data sharing. Therefore, we also show the performance of centralized systems that can obtain only a fraction of the data of the federated system, specifically 50% or 20% of the data; in other words, the federation is composed of sites that in total have double or five times the amount of data that can be collected or aggregated centrally. Table S1 shows the number of MRI scans per age bucket used to train the centralized (100%, 50%, 20%) and federated models. All models are evaluated against the same test set of 2,090 samples. Federated training significantly outperforms a centralized system when the federation reaches five times more data, which is feasible, and outperforms or matches centralized training when the federation reaches double the data, which is a reasonable assumption. We expect that without the burden of data sharing, much larger federated consortia can be formed and yield better analyses.

For the AD detection task, we analyzed three well-known studies: the Alzheimer's Disease Neuroimaging Initiative (ADNI),[37] with its three phases ADNI 1, ADNI 2, and ADNI 3; the Open Access Series of Imaging Studies (OASIS)[38]; and the Australian Imaging, Biomarkers & Lifestyle Flagship Study of Ageing (AIBL).[39] These studies contain T1-weighted brain MRIs taken from patients with different degrees of dementia and healthy subjects acting as controls. For our work, we used only images from control subjects and patients diagnosed with AD. These studies are longitudinal. To obtain unbiased performance estimates, all the samples for a given subject appeared either in the training or the test set. Table 1 shows the numbers of training and testing samples from each study and the target labels.

We compared the performance of models trained on each single study (ADNI-{1,2,3}, OASIS-3, and AIBL), a centralized model with the data of all studies, and models trained on a federation with an increasing number of sites/cohorts, that is, a federation of three sites, with each of the three ADNI phases (ADNI-{1,2,3}); a four-site federation (ADNI-{1,2,3} + OASIS-3); and a five-site federation (ADNI-{1,2,3} + OASIS-3 + AIBL). All environments trained the same 3D-CNN neural architecture (shown in Figure S3). In Table 2, we report every model's mean and standard deviation values over three execution runs. The model obtained by the federation outperforms models trained at any single site and has comparable AUC ROC with respect to the centralized model trained over all the data. The AUC ROC

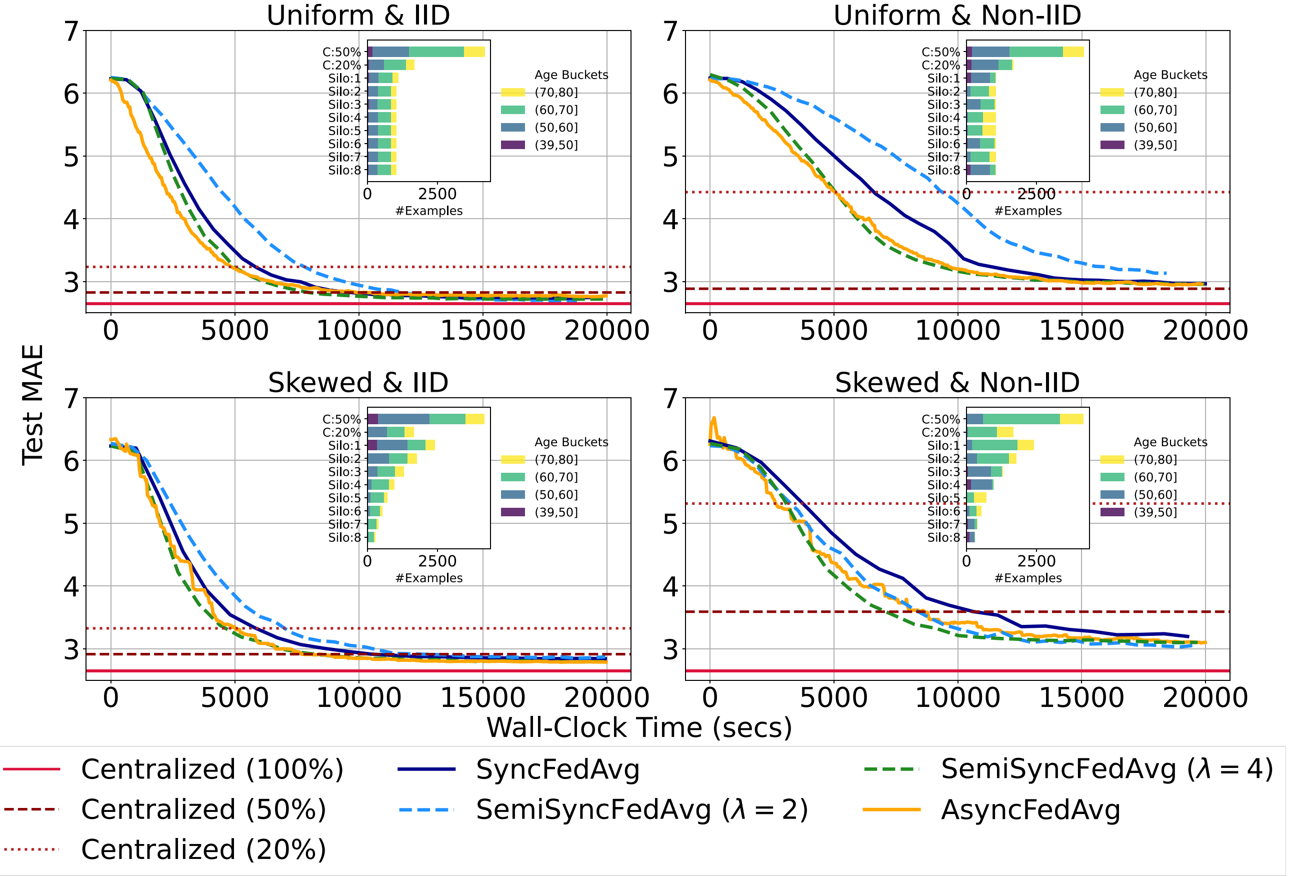


Figure 2. Comparing centralized and federated training on the BrainAGE task

Centralized (100%), (50%), and (25%) represent the models trained using all, half, and a quarter of the available data in a federation, respectively. For the last two cases, the centralized data were assembled starting from the site with the largest number of data samples. The federated models were trained using the synchronous, semi-synchronous, and asynchronous training policies. Federated training is comparable to centralized in the uniform and IID environment. For the hardest environment, skewed and non-IID, a federation that accumulates twice the data of a centralized consortium yields a significantly better performing model.

provides a robust measure of classifier performance, since it does not depend on a specific threshold. A larger difference in precision and recall values reflects a sensitivity to the classification threshold likely due to the class imbalance. The greater the number of cohorts participating in the federation, the better the predictive performance of the learned model is.

Fully homomorphic encryption can efficiently protect federated learning against attackers outside the federation

Neural models can memorize training data and are susceptible to model inversion attacks[32,33] or membership inference attacks.[34,35] Thus, if the sites shared unprotected models with the federation controller, or the models were captured in transit, an attacker may obtain private information. To prevent such attacks against the local neural models, we use FHE, specifically the CKKS scheme.[31] The sites encrypt their local models before transmission and the federated controller aggregates the models in an encrypted space.

Previous work[40,41] has studied secure aggregation using masking[40] for cross-device settings with frequent dropouts or for simple non-weighted aggregations over large plaintext spaces.[41] However, these approaches implicitly trust the controller with the global model,[40,41] whereas our work adopts a threat model where the controller is untrusted. FHE is crucial in this setting, as it ensures that no information about the local or global models is revealed to the controller during the aggregation process. Our findings underscore that privacy attacks, such as membership inference attacks, pose a significant privacy risk (see Figure 5). While employing FHE introduces a moderate computational and communication overhead (see Figures 4 and S5), the enhanced security and privacy guarantees, including safeguarding against inference attacks by a malicious controller and protecting the consortium's intellectual property, are well worth its cost.

We evaluated the learning performance of CKKS FHE on the BrainAGE prediction task over four federated learning environments using a 3D-CNN model with 3 million parameters (Figure S3). Figure 4 demonstrates that the inclusion of FHE, thanks to our optimizations, introduces a minimal runtime overhead compared to training without encryption. Specifically, the execution time for synchronous federated averaging[15] (SyncFedAvg) with encryption shows only a slight increase (~7%) compared to its unencrypted counterpart. Furthermore, Figure S5 shows

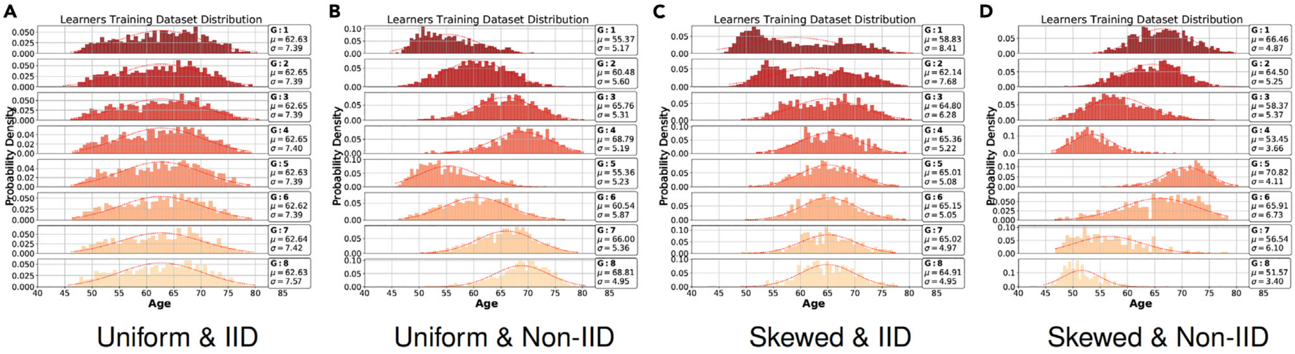


Figure 3. The age distribution of the UKBB MRI scans allocated to each silo/site/learner (denoted as G:*i*) on the four federated learning environments investigated in the BrainAGE prediction task. For every distribution the small white box next to the plot shows its standard deviation (σ) and mean (μ)

a model convergence comparison based on communication cost. These findings underscore the ability of our framework to offer robust privacy protections with only a marginal increase in training time, making it a practical solution for secure federated learning.

Federated learning is vulnerable to insider attacks

Our architecture uses HE for secure communication and aggregation of parameters. Thus, the system is protected against attacks from outside the federation or a compromised controller. However, each learner needs to decrypt the community model for local training. Therefore, a curious site inside the federation may attempt a model inversion attack[32,33] or a membership inference attack[34,35] against the community model (the local models of the other learners are protected through encryption). Model inversion attacks against the global federation model are impractical in realistic settings, since learners train the global model locally for a large number of local iterations, and the local models are aggregated. Therefore, the risk of leaking any identifiable information from the global model in realistic settings is limited.[42,43] In contrast, membership inference attacks are very successful. A site can use private local subject data to probe the federated community model and discover if such data were used for training. Discovering that the MRI scan of a particular subject was used for training the model can reveal information about a person's medical history or participation in some sensitive medical study.[44]

Figure 5 shows the increasing vulnerability of community models with each federation round. We used the features and architecture in Gupta et al.[44] to conduct the membership inference attacks. We measured vulnerability as the average accuracy of detecting that an MRI was used for training over 56 different datasets (see experimental procedures). As training progresses, the neural network learns more information about the samples, and it becomes easier to identify MRIs participating in training. Notably, in the uniform and IID environment, attack success reached 80%. Previous works[44–47] have also found a strong correlation between overfitting and vulnerability to these attacks. We see that data distribution across silos may impact vulnerability. Models trained over more homogeneous (IID) data distributions across silos are more vulnerable than heterogeneous data distributions (non-IID). Non-IID distributions often train slowly, with implicit regularization due to each local model being trained on different distributions, thus reducing overfitting. Overall, vulnerability increases during training, suggesting a trade-off between learning performance and privacy risk.

Federated training with gradient noise protects against membership attacks from insiders

The success of privacy attacks is often attributed to the ability of learning algorithms to memorize information about a single sample.[48] Therefore, defending against data privacy leakages involves limiting the information the learning algorithm may extract about each sample or limiting information in the neural network's weights. We explore approaches to limit the vulnerability to membership inference attacks: differential private training via DP-SGD[49] and stochastic gradient descent (SGD) with non-unique gradients.

Figure 6 shows the privacy and learning performance trade-off when sites are trained with small-magnitude Gaussian noise and our non-unique gradient approach. Both approaches can reduce the vulnerability of the global model to privacy attacks. The standard deviation of Gaussian noise used in our experiments is varied between 10^{-4} and 10^{-1}, which corresponds to $\varepsilon > 10^{10}$ for any reasonable δ, when considering (ε, δ)-DP. As such, when trained at such noise levels, the models would not have privacy guarantees of differential privacy. Nonetheless, even though the magnitude of Gaussian noise is much smaller than the theoretically required differential privacy guarantees, it effectively reduces membership inference attacks.

Varying capabilities across federated learning frameworks

Many federated learning architectures have recently become available to address optimization and system challenges for different domains, e.g., biomedical and finance, including OpenFL,[50] Nvidia FLARE,[51] Flower,[52] FedML,[53] IBM FL,[54] PriMIA,[55] Substra,[56] Fed-BioMed,[57] FATE,[58] FedScale,[59] and COINSTAC[60] frameworks. In Table 3, we provide a qualitative comparison for all the aforementioned federated learning frameworks from the perspective of supported functionalities, following the taxonomy in the works of Li et al.,[61] Kairouz

Table 1. Train and test data splits per cohort for the Alzheimer's disease prediction task

	Train set		Test set	
Cohort	Alzheimer's	controls	Alzheimer's	controls
OASIS-3	315	1,113	68	209
AIBL	113	642	28	160
ADNI-1	759	940	299	256
ADNI-2	534	1,137	185	399
ADNI-3	90	388	26	118
Total	1,811	4,220	606	1,142

et al.,[62] and Liu et al.[63] Even though COINSTAC provides a powerful platform for decentralized neuroimaging analysis, we did not consider it in our evaluation, since it has limited support for deep learning methods and is not tailored for federated learning settings. We compared the frameworks based on their functionalities offered by their open-source versions. Specifically, we considered six categories: deployment, data partitioning, security and privacy, topology, communication, and software.

In the deployment category, we considered the cases where a federated learning framework can execute a federated learning application in a simulated environment, i.e., as parallel processes/threads within a single server, distributedly across multiple servers (cross-silo) or edge devices (cross-device). All frameworks support simulation and cross-silo deployments. Flower, FedML, FATE, and FedScale have tailored capabilities for cross-device environments. The other frameworks could be applied in cross-device environments but are not a specific focus.

In the data partitioning category, we evaluated whether a system supports federated model training over federated learning environments with horizontally or vertically partitioned data distributions.[16] The horizontally partitioned learning environments are supported by all frameworks. Only FATE and FedML support machine and deep learning methods over the more challenging vertically partitioned data.[64] Given that data about a patient are often distributed across many health-care organizations, our immediate plan is to extend MetisFL to support vertical federated settings as well.

In the security and privacy category, we assessed whether a framework supports the TLS (Transport Layer Security) protocol, the type of the secure aggregation (SecAgg) protocol and the cryptographic library (Crypto lib) used for executing the protocol, and support model training under differential privacy (e.g., DP-SGD or local or central differential privacy[65]). All frameworks support private model training and execution using TLS; PriMIA is the only framework not supporting TLS. Regarding the secure aggregation protocol, Nvidia FLARE, IBM FL, FedML,[66] and MetisFL support FHE through the CKKS[31] construction scheme. FedML also uses a mask-based encryption approach through a native library implementation (see LightSecAgg[67]), similar to Flower (see Salvia[68]). PriMIA supports secure multiparty computation (SMPC) through the SPDZ[69] scheme. FATE supports SMPC and partial HE (PHE) through the Paillier (BatchCrypt[70]) construction scheme, and OpenFL operates on a hardware-integrated trusted execution environment (TEE[71]). Fed-BioMed uses a multiparty computation (MPC) protocol. Unfortunately, for Substra and FedScale, we could not find any information on their online documentation or open-source code on their secure aggregation schemes.

The topology category compares the federated learning topologies under which each system operates. Centralized topology refers to federated learning environments where communication across learners is established through a centralized controller. In contrast, decentralized topology refers to the federated learning environments where no controller is present, and learners communicate through a peer-to-peer protocol. Finally, hierarchical topology captures federated learning environments that may consist of a lead and a sub-controller.[19] Our analysis shows that all systems can operate in a centralized federated learning environment. FATE and FedML are the only frameworks supporting decentralized environments. None of the systems support hierarchical training.

Another category we considered in our evaluation is the communication protocol and the communication layer that each framework can support during federated execution. Even

Table 2. Alzheimer's disease prediction

Model	Accuracy	Precision	Recall	F1	AUC PR	AUC ROC
(C) ADNI-{1,2,3}	0.8570 ± 0.0090	0.7940 ± 0.0311	0.8270 ± 0.0288	0.8095 ± 0.0080	0.8639 ± 0.0052	0.8954 ± 0.0057
(C) OASIS-3	0.4428 ± 0.0194	0.3686 ± 0.0036	0.7447 ± 0.0518	0.4927 ± 0.0091	0.3396 ± 0.0020	0.4631 ± 0.0047
(C) AIBL	0.8050 ± 0.0044	0.7246 ± 0.0153	0.7577 ± 0.0172	0.7405 ± 0.0022	0.7990 ± 0.0005	0.8526 ± 0.0017
(C) ADNI-{1,2,3} + OASIS-3 + AIBL*	0.8612 ± 0.0106*	0.7977 ± 0.0350*	0.8287 ± 0.0271*	0.8122 ± 0.0091*	0.8683 ± 0.0130*	0.8986 ± 0.0051*
(F) ADNI-{1,2,3}	0.8462 ± 0.0043	0.8148 ± 0.0189	0.8048 ± 0.0194	0.8095 ± 0.0038	0.8791 ± 0.0039	0.8967 ± 0.0012
(F) ADNI-{1,2,3} + OASIS-3	0.8474 ± 0.0073	0.7955 ± 0.0126	0.8296 ± 0.0026	0.8121 ± 0.0077	0.8766 ± 0.0007	0.8920 ± 0.0014
(F) ADNI-{1,2,3} + OASIS-3 + AIBL*	0.8633 ± 0.0013*	0.8098 ± 0.0043*	0.8132 ± 0.0097*	0.8114 ± 0.0031*	0.8682 ± 0.0009*	0.8971 ± 0.0006*

Test results on a global stratified test dataset (five sites), and for each dataset by itself in a centralized environment (ADNI-{1, 2, 3}, OASIS-3, and AIBL), and a federation of three sites (ADNI-{1,2,3}), four sites (ADNI-{1,2,3} + OASIS-3), and all five sites (ADNI-{1,2,3} + OASIS-3 + AIBL). We denote the centralized and federated environments with (C) and (F). In the federated environments, each dataset was located at a different site. Centralized environments were trained over all the corresponding datasets. The evaluation was conducted over three different runs. Values with asterisks (*) represent the best performance for the centralized (C) and federated environments (F).

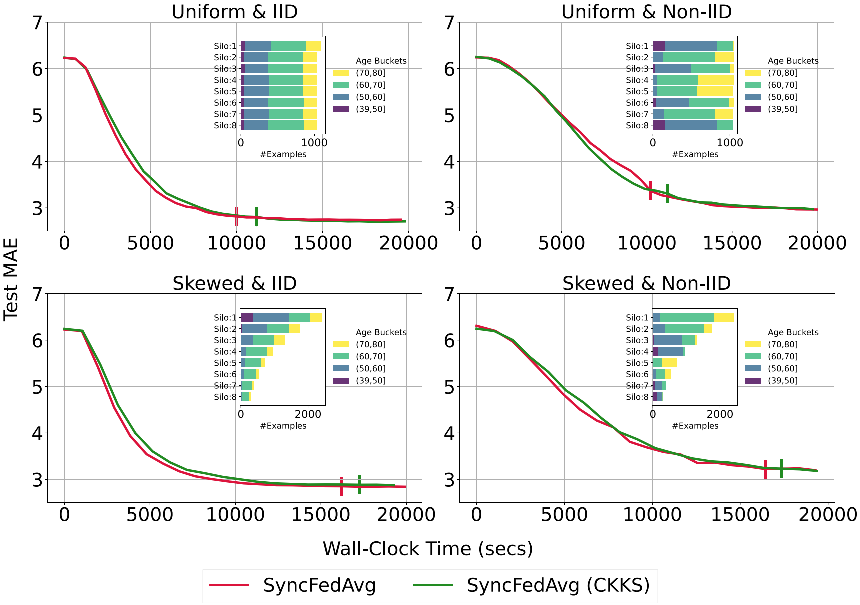


Figure 4. Synchronous federated learning training (SyncFedAvg) with and without homomorphic encryption using the CKKS scheme on the BrainAGE task for the 3D-CNN model

The vertical markers represent the training time it takes for each approach to complete 20 federation rounds. Training with encryption incurs only 7% runtime overhead. SyncFedAvg, synchronous federated averaging.

though all systems support synchronous communication and aggregation, they lack support for asynchronous protocols. In contrast, MetisFL supports synchronous (including semi-synchronous) and asynchronous execution. According to Flower's and FedML's documentation, they plan to support asynchronous execution soon. Concerning the communication layer, every reported system uses gRPC to establish communication across all system services, except for FedML, IBM FL, and Fed-BioMed, which also use gRPC, MPI, and MQTT protocols; AMQP; and MQTT, respectively.

Compared to the previously reported dimensions,[61–63] we also considered the programming language used to develop each component in the federated learning system, namely the controller and the learner. Across all reported systems, both components are developed in Python, except for Substra and MetisFL. Substra's controller is implemented in Go, while MetisFL's aggregator is implemented in C++. In our original implementation,[72] we also developed the aggregator component in Python. However, this approach led to a high latency when aggregating large-sized models and scheduling training and evaluation jobs across multiple clients (e.g., >200) due to Python's limited memory management capabilities. Python internally relies on the global interpreter lock (GIL) for proper thread management, and this hindered the concurrent execution of tasks and, therefore, slowed down the federated execution workflow dramatically.

In MetisFL, the federation controller is considered a first-class citizen of the entire system. By redesigning and refactoring the controller in a native C++ implementation, we were able to significantly optimize a range of controller-specific operations (e.g., tensor aggregation and network transmission) and achieve a 10- to 100-fold improvement against other leading FL frameworks.[30] Figure 7 demonstrates an evaluation of MetisFL compared to existing open-source frameworks regarding the total federation time, including model training, aggregation, and evaluation. The evaluation is conducted with a multilayer perceptron (MLP) model consisting of 10 million trainable parameters over an increasing number of learners (10, 25, 20, 100, 100, or 200), with full client participation at every round. As shown in the figure, for large-scale federations consisting of 100 and 200 learners, MetisFL with and without the use of OpenMP[73] leads to a 10-fold improvement compared to Flower and FedML, while other frameworks, such as IBM FL and NVFlare, fail to complete the federation round within a reasonable time. A more thorough and fine-grained evaluation of other system metrics can be found in our recent work.[30]

DISCUSSION

Dementia affects more than 50 million people worldwide, and this number could exceed 152 million by 2050,[74] with AD being the leading cause. Recently, deep learning has been applied to identify AD from structural brain MRI scans.[10,75–77] Similar in spirit, the BrainAGE task is another pathway toward assessing the acceleration or deceleration of an individual's brain aging through structural MRI scans. The difference between the true chronological age and the predicted age of the brain is considered an important biomarker for early detection of age-associated neurodegenerative and neuropsychiatric diseases,[78,79] such as cognitive impairements,[80] schizophrenia,[81] and chronic pain.[82]

Recently, deep learning methods based on RNN[83,84] and CNN[85–88] architectures have demonstrated accurate brain age predictions. We extended these methods to train a 3D-CNN model over centralized and federated environments with highly heterogeneous data distributions[72,89] for the BrainAGE and AD prediction tasks. We show that federated learning can achieve comparable performance relative to centralized training. We posit that, ultimately, federated consortia will allow one to analyze much greater quantities of data, since they sidestep many of the challenges of centralized data sharing.

We use HE[90] to ensure that models are protected from attacks from outside the federation. HE is a public-key encryption scheme[91] that enables certain computations (e.g., additions or multiplications) to be directly performed over encrypted data without decrypting them first. This distinct computational property renders HE a valuable cryptographic scheme for preserving data privacy in distributed settings, as untrusted parties can be tasked with performing operations over ciphertexts. In our setting, we consider an honest-but-curious threat model and assume the participating parties do not collude with each other. To ensure secure model communication and

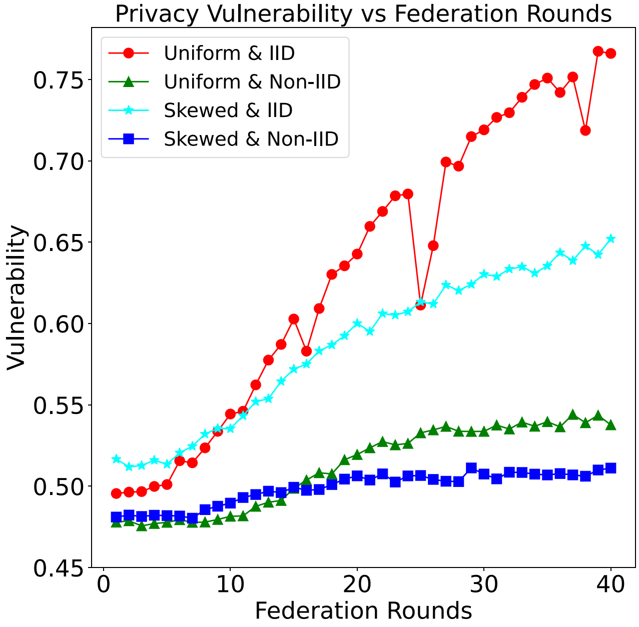


Figure 5. Privacy vulnerability increases with federation rounds

Vulnerability is the average accuracy of distinguishing train samples vs. unseen samples across sites.

aggregation, we use CKKS, an FHE scheme.[31] Compared to the Paillier scheme used in past works,[70,92] CKKS supports arithmetic operations over real and complex data and is orders of magnitudes faster and can support an unbounded amount of additions and multiplications over encrypted data. While some previous works[70,92,93] may leak the global model to the controller, our protocol ensures that no global model leakage occurs at the controller.

Even though federated learning avoids sharing datasets, and model transmission and aggregation can be protected through encryption, the clients have access to the unencrypted model, whose parameters may reveal private information. Various works have highlighted this through practical privacy attacks such as model inversion[32,33] and membership inference[34,35] attacks, in both centralized and federated settings. Researchers have focused on reducing overfitting,[46,47] limiting information in activations and weights,[48] or using differential private mechanisms to alleviate such privacy concerns. Learning under differential privacy is particularly attractive, as it comes with theoretically solid worst-case guarantees. However, these works assume different threat models. For example, Wei et al.[94] assume that the server can be trusted, whereas Noble et al.[95] and Zhao et al.[96] consider a stricter threat model, considering the server as honest but curious, similar to us. Rather than using a theoretical upper-bound measure of privacy, we focus on a more practical measure (i.e., membership inference attacks). We study membership inference attacks in our framework using the white-box experimental setup from Gupta et al.[44] We assume that the attacker has access to the model, some samples that participated in the training, and some samples the attacker is curious about. We show that federated training with noise protects the models from attacks and the trade-off between protection and learning performance.

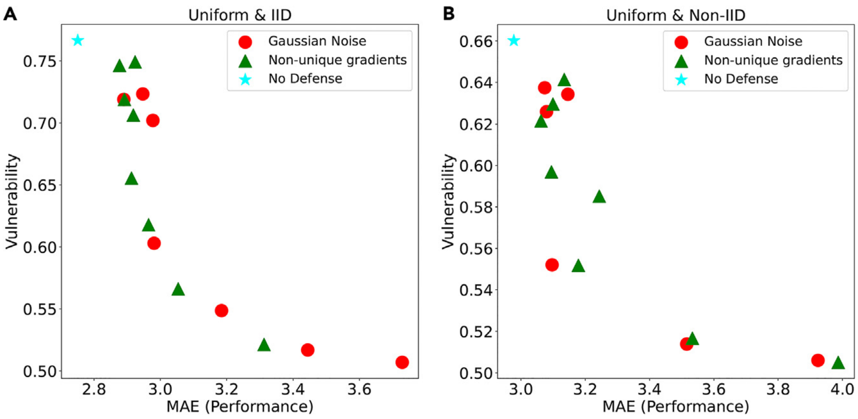
In summary, we have presented the MetisFL secure and private federated learning system, a practical, extensible architecture supporting a variety of communication protocols with strong privacy and security mechanisms. Specifically, MetisFL protects against attacks from outside the federation through efficient HE and against insider attacks by adding small, targeted noise during federated training. We demonstrated its efficacy in neuroimaging tasks, BrainAGE estimation, and AD prediction over challenging statistically heterogeneous environments. We showed that federated learning can achieve the same learning performance as centralized learning in realistic environments. In hard heterogeneous environments, a small performance gap remains. We expect that centralized consortia, which require data sharing, will include fewer sites than federated consortia, which do not share data. We posit that the larger consortia using federated learning promise to yield better analysis and greater advances in biomedical research.

EXPERIMENTAL PROCEDURES

Resource availability

Lead contact

Correspondence and requests for materials should be addressed to Dimitris Stripelis at stripeli@isi.edu.



Uniform & IID

Uniform & Non-IID

Figure 6. Vulnerability vs. performance trade-off when training sites with differential privacy (Gaussian noise) and non-unique gradients approaches

Lower vulnerability and lower MAE are desired, i.e., points toward the bottom left are better.

Table 3. A qualitative comparison of different federated learning systems

Category	OpenFL	NVFlare	Flower	FedML	IBM FL	PriMIA	Substra	Fed-BioMed	FATE	FedScale	MetisFL
Deployment											
Simulation	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Cross-silo	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Cross-device	+	+	✓	✓	+	+	×	+	✓	✓	+
Data partitioning											
Horizontal	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Vertical	×	×	×	✓	×	×	×	×	✓	×	×
Security and privacy											
Private training	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
SecAgg	TEE	FHE	masking	masking | FHE	FHE	SMPC	?	MPC	SMPC | PHE	?	FHE
Crypto lib	Graphene	TenSeal	native	native | PALISADE	HElayers	PySyft	?	MP-SPDZ	native	?	PALISADE
TLS	✓	✓	✓	✓	✓	×	✓	✓	✓	✓	✓
Topology											
Centralized	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Decentralized	×	×	×	✓	×	×	✓	×	✓	×	×
Hierarchical	×	×	×	×	×	×	×	×	×	×	×
Communication											
Synchronous	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Asynchronous	×	×	×	×	×	×	×	×	×	×	✓
Network	gRPC	gRPC	gRPC	gRPC | MPI | MQTT	AMQP	gRPC	gRPC	MQTT	gRPC	gRPC	gRPC
Software											
Learner	Python	Python	Python	Python	Python	Python	Python	Python	Python	Python	Python
Controller	Python	Python	Python	Python	Python	Python	Go	Python	Python	Python	C++

✓, supported; ×, not supported; +, possible, but not focus; ?, unknown.

Materials availability

This study did not generate materials.

Data and code availability

This is a study of previously collected, anonymized, deidentified data available in a public repository. For the UKBB data, access was approved by the UKBB under application no. 11559. All other datasets analyzed during the current study are available in the ADNI (https://adni.loni.usc.edu/), OASIS-3 (https://www.oasis-brains.org/), and AIBL (https://aibl.csiro.au/) repositories. Further information is available in published articles for ADNI[97] and OASIS-3.[38] All the original code of the MetisFL framework has been deposited at Zenodo[29] (https://dx.doi.org/10.5281/zenodo.11411754) under https://doi.org/10.5281/zenodo.11411754.1 and is publicly available as of the date of publication. Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

Federated learning

A federated learning environment consists of N sites (learners or clients) that jointly train a machine learning model, often a neural network. The goal is to find the model parameters w^* that optimize the global objective function $F(w): w^* = \underset{w}{\text{argmin}}\, F(w) = \sum_{k=1}^{N} \frac{p_k}{P} F_k(w)$, where $F_k(w)$ denotes the local objective function of learner $k \in N$ optimized over its local training dataset D_k. The global model, $F(w)$, is computed as a weighted average of the learners' local models, $P = \sum_k^N |p_k|$. A typical policy,[15] which we follow in this paper, is to weigh each local model based on the number of training examples it was trained on, i.e., $p_k = |D_k|$, $P = \sum_k^N |D_k|$, although other methods are possible.[16,98,99] Typically, each learner uses SGD to optimize its local objective on its local dataset. At a given synchronization point, each learner shares its local model parameters with the federation controller, which aggregates the local models (e.g., using weighted average) to compute a global (or community) model and sends it back to the learners, and the process repeats. Each such cycle is called a federation round. This iterative process was first introduced in the seminal work of Tian and Gu,[15] and it is termed as FedAvg.

More recent works[100–102] have proposed a more general federated learning optimization framework where the optimization problem is split into server side (global) and client side (local). Server side aims to optimize the merging/aggregation rule of the learners' local model updates, and the client side aims to optimize learners' local model training. During training, a learner shares local model parameters only with the federation controller. Each local dataset remains private.

Predicting BrainAGE

In our experiments we use the same 3D-CNN architecture as in Stripelis et al.,[89] but without the dropout layer. This slight modification improved the learning performance on the BrainAGE task for both the centralized and the federated models across all environments. The training and testing datasets follow Stripelis et al.[72,89] We selected 10,446 scans from the UKBB[36] dataset with no indication of neurological pathology and no psychiatric diagnosis as defined by the ICD-10 criteria. All scans were evaluated with a manual quality control procedure and processed using a standard preprocessing pipeline with non-parametric intensity normalization for bias field correction and brain extraction using FreeSurfer and linear registration to a $(2\text{ mm})^3$ UKBB minimum deformation template using FSL FLIRT. The final dimensions of each scan were 91 × 109 × 91. Of the 10,446 scans, we used

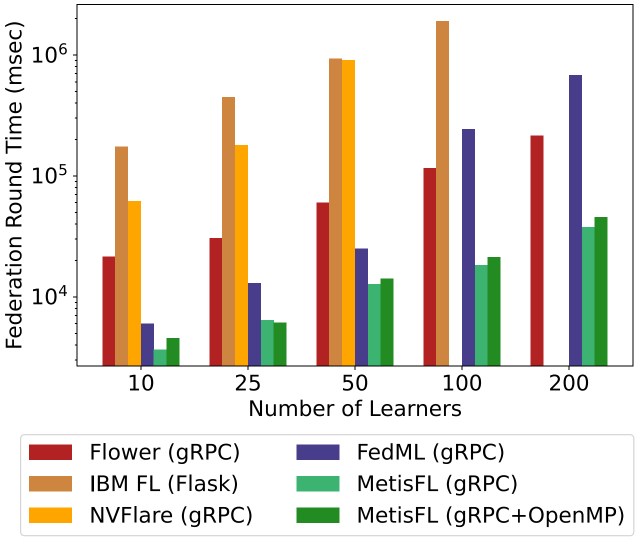


Figure 7. A comparison of the MetisFL framework to other open-source federated learning frameworks: Flower, IBM FL, NVFlare, and FedML

Inside the parentheses, we state the communication protocol (gRPC or Flask) and the optimizations (OpenMP) used to run each framework. The performance across all frameworks is measured as the total time (in seconds, logarithmic y axis) it takes to complete a federation round when training an MLP model with 10 million parameters over an increasing number of learners, i.e., 10, 25, 50, 100, or 200. MetisFL greatly improves federated training time, by almost a 10-fold improvement compared to Flower and IBM FL, while NVFlare and FedML fail to complete within a reasonable time in large-scale federations, i.e., federations with 100 and 200 learners.

8,356 for training and 2,090 for testing. We generated four computationally and statistically heterogeneous federated learning environments comprising eight sites (learners). Computationally, the first four learners were equipped with NVIDIA GeForce GTX 1080 Ti GPUs, while the last four had (faster) Quadro RTX 8000 GPUs.

For data amounts, we considered both uniform, an equal number of training samples per learner, and skewed, a decreasing amount of training samples for each learner. For data distributions, we considered both IID, in which the local data distribution of each learner contains scans with the same distribution as the global age distribution, and non-IID, different age distributions. We tested synchronous (SyncFedAvg), asynchronous (AsyncFedAvg), and semi-synchronous (SemiSyncFedAvg)[72] federated training policies, all using the training data size as a weighting rule. We evaluated the performance of all policies against the same holdout test dataset to estimate the MAE between individuals' true chronological brain age and their predicted age.

Each site (learner) is trained using SGD with a learning rate of 5×10^{-5} and a batch size of 1. For SyncFedAvg and AsyncFedAvg, local training was performed over four epochs. For SemiSyncFedAvg, we evaluated synchronization periods λ equivalent to the time it took the slowest learner to complete four or two epochs. AsyncFedAvg uses the caching method we introduced in Stripelis et al.,[72] significantly improving performance. The controller stores each learner's most recently committed local model in a cache. When a learner issues an update request, the controller replaces its previously cached model and computes the new global model by performing a weighted average using all cached models.

Figure 2 compares the learning convergence of the training policies based on elapsed execution (wall-clock) time; we also provide a comparison in terms of communication cost in the supplemental information. In IID environments, both for uniform and for skewed data amounts, federated training achieves comparable learning performance (MAE) relative to centralized training. The asynchronous protocol, AsyncFedAvg, is competitive in task performance but requires significantly more communication. SemiSync has fast convergence with low communication costs.

With centralized 20%, 50%, and 100%, we aim to emulate small, medium, and large research consortia that have established data sharing agreements to share their local data with a central authority for further analysis. To generate the centralized datasets for every environment, we start assigning data samples from the silo owning the majority of data samples (i.e., silo:1) until we reach 20% or 50% of the total data. That is, we assume that the first few (largest) sites in the consortia decide to share data. All models (centralized and federated) are evaluated on the same test dataset.

Predicting AD

In our evaluation, we studied three prominent AD studies: the three phases of ADNI,[37] OASIS-3,[38] and AIBL.[39] Images across all sites were preprocessed following the pipeline in Dhinagar et al..[75] First, images were reoriented using fslreorient2std (FSL v.6.0.1) so as to match the orientation of standard template images. Then, brain extraction was performed: skull parts in the image were removed using the HD-BET CPU implementation, and gray- and white-matter masks were extracted using FSL-FAST (FSL v.6.0.1 Automated Segmentation tool). An intensity normalization step (N4 bias field correction) using ANTs (v.2.2) followed. Next, linear registration to a UKBB minimum deformation template was obtained by using the FSL-FLIRT (FSL v.6.0.1 Linear Image Registration tool) with 6 degrees of freedom. Finally, an isometric voxel resampling to 2 mm was applied using the ANTs ResampleImage tool (v.2.2). The actual size of the images after the preprocessing was volumes of 91 × 109 × 91 voxels. In our previous work,[103,104] we developed methods for scanner invariant representations and imaging harmonization; however, we did not apply these techniques in the experiments, since the images had already been processed through a common pipeline.

We trained a 3D-CNN neural model over a federation of three (ADNI phases), four (ADNI phases + OASIS), and five learners (ADNI phases + OASIS + AIBL). Table 2 shows the performance of the federated and the centralized models. The federated models were trained using the synchronous protocol (i.e., SyncFedAvg) for 40 federation rounds, with each learner training locally for four epochs in-between rounds. The centralized models were trained for 100 epochs. All models were trained using Adam with Weight Decay with a learning rate of 1e−5 and weight decay of 1e−4. All experiments were run three times, and the results show the average and standard deviation of the metrics.

Secure FL using FHE

Figure 8 presents the secure federated training pipeline of our MetisFL system. We use HE to communicate the (encrypted) local and global models between the federation controller and the learners and compute the new global model by aggregating learners' local models in an encrypted space. Training starts with an initial configuration phase, where the federation driver generates the HE key pair (private and public key) and the original neural model state. The federation controller receives only the model definition and the public key from the driver, since it only needs to perform the private weighted aggregation of local models. In contrast, the learners need private and public keys during training. The private key is used to decrypt the encrypted global model received by the controller to perform their local training (or model evaluation) over their local private dataset, and the public key is used to encrypt the locally trained model before being shared with the controller.

We used a similar training pipeline in Stripelis et al.[26] However, in our previous work, we encrypted the entire model into a single ciphertext, which created scalability issues for large models. To mitigate this, in MetisFL, we encrypted the model on a matrix-by-matrix basis, allowing for a collection of ciphertexts to be communicated between learners and the controller instead of just a single ciphertext. Thereafter, the controller performs the private weighted aggregation over the collection of ciphertexts from all learners. In addition, we divided the model parameters into batches processed in parallel, leading to a much faster encrypted computation.

Figure 9 demonstrates the effect of batching multiple model parameters on the size of the encrypted model. Batching multiple parameters into a single ciphertext helps us reduce the overall model size and allows us to leverage SIMD (single instruction multiple data) for faster processing of encrypted

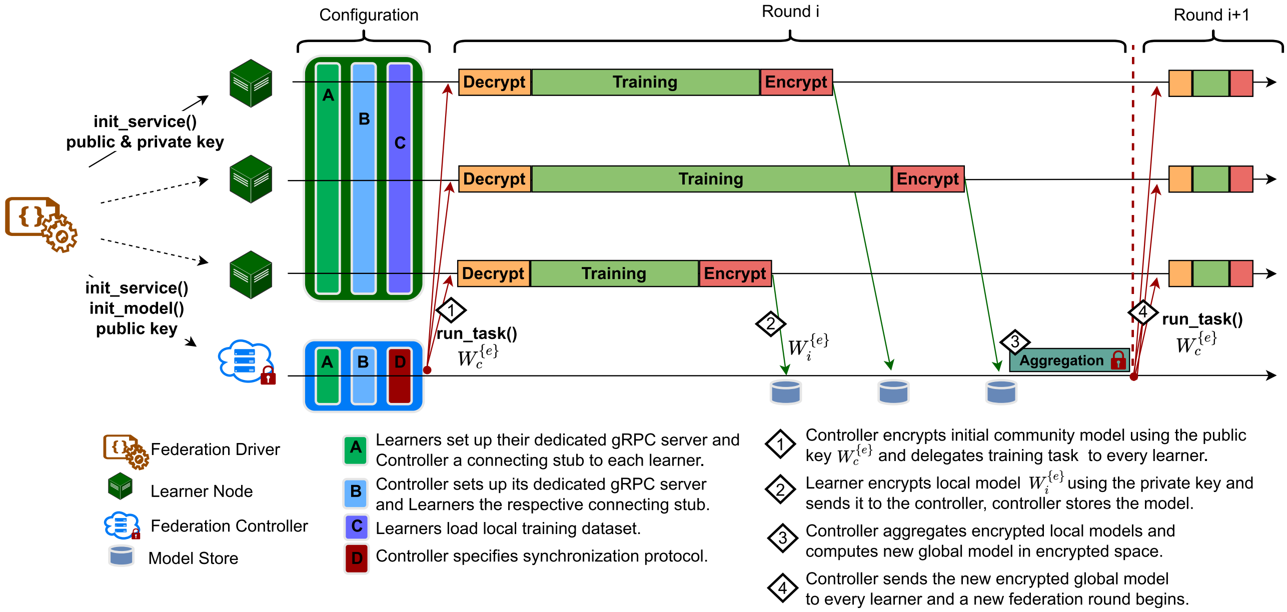

Configuration
Round i
Round i+1
init_service()
public & private key
init_service()
init_model()
public key
A
B
C
D
Decrypt
Training
Encrypt
run_task()
$W_c^{\{e\}}$
$W_i^{\{e\}}$
Aggregation
1
2
3
4
Federation Driver
Learner Node
Federation Controller
Model Store
A Learners set up their dedicated gRPC server and Controller a connecting stub to each learner.
B Controller sets up its dedicated gRPC server and Learners the respective connecting stub.
C Learners load local training dataset.
D Controller specifies synchronization protocol.
1 Controller encrypts initial community model using the public key $W_c^{\{e\}}$ and delegates training task to every learner.
2 Learner encrypts local model $W_i^{\{e\}}$ using the private key and sends it to the controller, controller stores the model.
3 Controller aggregates encrypted local models and computes new global model in encrypted space.
4 Controller sends the new encrypted global model to every learner and a new federation round begins.

Figure 8. Homomorphic encryption training pipeline in the MetisFL framework using the synchronous communication protocol
After the federation configuration, the controller sends the original encrypted model to each learner, the learners decrypt and train the received global model and then encrypt and send the new local model to the controller, and the controller aggregates the encrypted models, and a new federation round begins.

models. The CKKS parameters are multiplicative depth of 1, 52 scale factor bits, batch size of 4,096, and security level of 128 bits.

Membership inference attacks

To conduct the membership inference attacks, we used the same features and architecture as in Gupta et al.[44] for training models to predict membership. We trained the attack model for each learner by creating a training set from their training samples and samples not seen during the model training. Finally, we computed how accurately we can predict the training samples of all the other learners vs. samples not used for training. We created a balanced set of samples used for training and unseen samples (i.e., the test samples) from each learner to compute this accuracy. We report the average of these 56 accuracy values as the vulnerability score—each learner trains attack models (i.e., eight attackers) and predicts train vs. unseen on samples from the seven other learners.

Gaussian noise

Differential privacy is a formal framework to reason about privacy. A differential private training mechanism ensures that the outcomes (i.e., final model weights) do not change much between two training sets that differ by one example.[105] For training brain age models with differential privacy guarantees, we use the DP-SGD algorithm.[49] Briefly, the principal modifications to SGD to limit the influence of a single sample are to clip the gradients from each sample not to exceed a maximum norm and to add spherical Gaussian noise. We update each learner during federated training with these private gradients. During initial experiments, we found that achieving non-vacuous differential privacy guarantees requires adding significant Gaussian noise to the gradients, which annihilates learning performance. However, we observed that practical privacy attacks, such as membership inference attacks, can be thwarted by adding Gaussian noise of much smaller

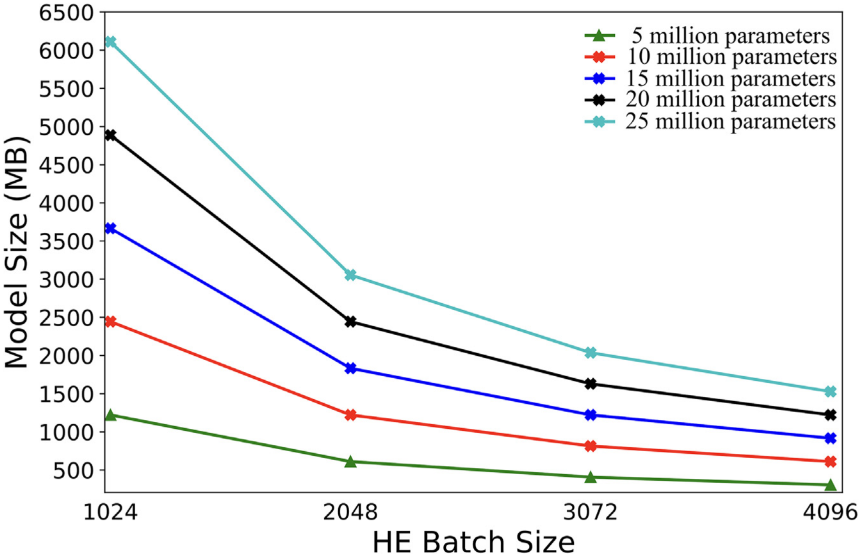


Figure 9. Relation between the encrypted model size and the batch size of the CKKS scheme
The encrypted model size is reduced as we increase the batch size of the FHE scheme (scaling factor bits, 52; security level, 128).

magnitudes.[106] Therefore, we evaluated training with gradients with a small additive Gaussian noise.

Learning with non-unique gradients

To learn good machine learning models, we would like to extract patterns while ignoring information about specific samples. Training models using gradient descent can leak an individual's information during training because there is no restriction on what information a sample may contribute. Thus, the model may preserve information unique to each individual, leaking privacy. Differential privacy adds the same noise to all gradients to limit the information or influence of a single sample on the neural network, but that may also destroy useful information in an attempt to reduce memorization. We investigated removing unique information from each sample's gradient and training with only non-unique parts. We computed the gradient of the loss (L) with respect to parameters (θ) for each sample (x_i, y_i) in a batch (B), i.e., $g_i = \nabla_\theta L(f(x_i;\theta), y_i)\ \forall i \in \{1 \ldots B\}$. To compute the non-unique part, we projected each gradient vector on the subspace spanned by the rest of the gradient vectors (g_i^{span}). We considered the residual part as the unique information about each sample (i.e., $g_i^{unique} = g_i - g_i^{span}$. Ideally, we would like to train with only the non-unique part. However, we observed that it may harm the performance too much. Therefore, we downweighed the effect of the unique part and used $\widehat{g}_i = g_i^{span} + \alpha g_i^{unique}, \alpha < 1$ to update the model at local learners; α is a hyperparameter that we tune to trade off privacy and performance.

We hypothesized that the small additive noise is enough to reduce the mutual information between data and neural network outputs/activations, which limits the success of membership inference attacks.[48] In IID environments, non-unique gradients perform similar to adding Gaussian noise. However, they are significantly faster to train. The Gaussian noise models required training for 40 rounds, whereas the non-unique gradients required only 25 rounds. Training with Gaussian noise in non-IID environments provides a better trade-off than non-unique gradients. This may be due to learners overfitting to private datasets earlier in training, thus deviating from the community model. In summary, both small-magnitude Gaussian noise added to gradients and non-unique gradients are effective at preventing membership attacks.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.patter.2024.101031.

ACKNOWLEDGMENTS

This research was supported in part by the Defense Advanced Research Projects Agency (DARPA) under contract HR00112090104 and in part by the National Institutes of Health (NIH) under grants U01AG068057 and RF1AG051710. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of DARPA, NIH, or the US government.

AUTHOR CONTRIBUTIONS

D.S. co-led the effort, developed the system architecture and experiments, and analyzed findings. U.G. developed the privacy components, co-developed the unique gradients method, and analyzed the findings. N.D. defined the deep learning models used for the BrainAGE and AD experiments. H.S. and T.G. helped with the homomorphic encryption development. C.A. helped with the Metis federated learning system design and development. R.S. helped with the AD prediction experiments. G.V.S. advised on machine learning methods, led the data privacy analysis, and co-developed the unique gradients method. S.R. and M.N. advised on homomorphic encryption methods and system security development. P.M.T. advised on neuroimaging data analysis and experiments. J.L.A. led the effort, co-designed the federated learning architecture and training algorithms, and advised on experiment design and data analysis findings. Every author has approved the submitted version and has agreed both to be personally accountable for the author's own contributions and to ensure that questions related to the accuracy or integrity of any part of the work, even ones in which the author was not personally involved, are appropriately investigated and resolved and the resolution be documented in the literature.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: October 10, 2023
Revised: April 4, 2024
Accepted: June 6, 2024
Published: August 1, 2024

REFERENCES

1. Wainberg, M., Merico, D., Delong, A., and Frey, B.J. (2018). Deep learning in biomedicine. Nat. Biotechnol. *36*, 829–838. https://doi.org/10.1038/nbt.4233.
2. Suzuki, K. (2017). Overview of deep learning in medical imaging. Radiol. Phys. Technol. *10*, 257–273. https://doi.org/10.1007/s12194-017-0406-5.
3. Zhu, B., Liu, J.Z., Cauley, S.F., Rosen, B.R., and Rosen, M.S. (2018). Image reconstruction by domain-transform manifold learning. Nature *555*, 487–492. https://doi.org/10.1038/nature25988.
4. Dalca, A.V., Yu, E., Golland, P., Fischl, B., Sabuncu, M.R., and Eugenio Iglesias, J. (2019). Unsupervised deep learning for bayesian brain mri segmentation. In International Conference on Medical Image Computing and Computer-Assisted Intervention (Springer), pp. 356–365. https://doi.org/10.1007/978-3-030-32248-9_40.
5. Cho, J., Kim, Y.J., Sunwoo, L., Lee, G.P., Nguyen, T.Q., Cho, S.J., Baik, S.H., Bae, Y.J., Choi, B.S., Jung, C., et al. (2021). Deep learning-based computer-aided detection system for automated treatment response assessment of brain metastases on 3d mri. Front. Oncol. *11*, 739639. https://doi.org/10.3389/fonc.2021.739639.
6. Kofler, F., Berger, C., Waldmannstetter, D., Lipkova, J., Ezhov, I., Tetteh, G., Kirschke, J., Zimmer, C., Wiestler, B., and Menze, B.H. (2020). Brats toolkit: Translating brats brain tumor segmentation algorithms into clinical and scientific practice. Front. Neurosci. *14*, 125. https://doi.org/10.3389/fnins.2020.00125.
7. Aksman, L.M., Wijeratne, P.A., Oxtoby, N.P., Eshaghi, A., Shand, C., Altmann, A., Alexander, D.C., and Young, A.L. (2021). pysustain: A python implementation of the subtype and stage inference algorithm. SoftwareX *16*, 100811. https://doi.org/10.1016/j.softx.2021.100811.
8. Young, A.L., Vogel, J.W., Aksman, L.M., Wijeratne, P.A., Eshaghi, A., Oxtoby, N.P., Williams, S.C., and Alexander, D.C.; Alzheimer's Disease Neuroimaging Initiative (2021). Ordinal sustain: Subtype and stage inference for clinical scores, visual ratings, and other ordinal data. Front. Artif. Intell. *4*, 613261. https://doi.org/10.3389/frai.2021.613261.
9. Ezzati, A., Abdulkadir, A., Jack Jr, C.R., Thompson, P.M., Harvey, D.J., Truelove-Hill, M., Sreepada, L.P., Davatzikos, C., Initiative, A.D.N., and Lipton, R.B. (2021). Predictive value of atn biomarker profiles in estimating disease progression in alzheimer's disease dementia. Alzheimers Dement. *17*, 1855–1867. https://doi.org/10.1002/alz.12491.
10. Lu, B., Li, H.-X., Chang, Z.-K., Li, L., Chen, N.-X., Zhu, Z.-C., Zhou, H.-X., Li, X.-Y., Wang, Y.-W., Cui, S.-X., et al. (2022). A practical alzheimer's disease classifier via brain imaging-based deep learning on 85,721 samples. J. Big Data *9*, 101. https://doi.org/10.1186/s40537-022-00650-y.
11. Thompson, P.M., Jahanshad, N., Ching, C.R., Salminen, L.E., Thomopoulos, S.I., Bright, J., Baune, B.T., Bertolín, S., Bralten, J., Bruin, W.B., et al. (2020). Enigma and global neuroscience: A decade of large-scale studies of the brain in health and disease across more than 40 countries. Transl. Psychiatry *10*, 100. https://doi.org/10.1038/s41398-020-0705-1.
12. Bischoff-Grethe, A., Ozyurt, I.B., Busa, E., Quinn, B.T., Fennema-Notestine, C., Clark, C.P., Morris, S., Bondi, M.W., Jernigan, T.L., Dale, A.M., et al. (2007). A technique for the deidentification of structural brain MR images. Hum. Brain Mapp. *28*, 892–903. https://doi.org/10.1002/hbm.20312.

13. Schimke, N., and Hale, J. (2011). Quickshear defacing for neuroimages. In 2nd USENIX Workshop on Health Security and Privacy (HealthSec 11) (USENIX Association) https://www.usenix.org/conference/healthsec11/quickshear-defacing-neuroimages.

14. Milchenko, M., and Marcus, D. (2013). Obscuring Surface Anatomy in Volumetric Imaging Data. Neuroinformatics *11*, 65–75. https://doi.org/10.1007/s12021-012-9160-3.

15. Tian, L., and Gu, Q. (2017). Communication-efficient Distributed Sparse Linear Discriminant Analysis. In Proceedings of the 20th International Conference on Artificial Intelligence and Statistics vol. 54 of *Proceedings of Machine Learning Research*, A. Singh and J. Zhu, eds. (PMLR), pp. 1178–1187. https://proceedings.mlr.press/v54/tian17a.html.

16. Yang, Q., Liu, Y., Chen, T., and Tong, Y. (2019). Federated machine learning: Concept and applications. ACM Trans. Intell. Syst. Technol. *10*, 1–19. https://doi.org/10.1145/3298981.

17. Li, T., Sahu, A.K., Talwalkar, A., and Smith, V. (2020). Federated learning: Challenges, methods, and future directions. IEEE Signal Process. Mag. *37*, 50–60. https://doi.org/10.1007/978-3-030-85559-8_13.

18. Jordan, M.I., Lee, J.D., and Yang, Y. (2019). Communication-efficient distributed statistical inference. J. Am. Stat. Assoc. *114*, 668–681. https://doi.org/10.1080/01621459.2018.1429274.

19. Rieke, N., Hancox, J., Li, W., Milletari, F., Roth, H.R., Albarqouni, S., Bakas, S., Galtier, M.N., Landman, B.A., Maier-Hein, K., et al. (2020). The future of digital health with federated learning. NPJ Digit. Med. *3*, 1–7. https://doi.org/10.1038/s41746-020-00323-1.

20. Lee, J., Sun, J., Wang, F., Wang, S., Jun, C.-H., and Jiang, X. (2018). Privacy-preserving patient similarity learning in a federated environment: development and analysis. JMIR Med. Inform. *6*, 7744. https://doi.org/10.2196/medinform.7744.

21. Sheller, M.J., Reina, G.A., Edwards, B., Martin, J., and Bakas, S. (2018). Multi-institutional deep learning modeling without sharing patient data: A feasibility study on brain tumor segmentation. In International MICCAI Brainlesion Workshop (Springer), pp. 92–104. https://doi.org/10.1007/978-3-030-11723-8_9.

22. Silva, S., Gutman, B.A., Romero, E., Thompson, P.M., Altmann, A., and Lorenzi, M. (2019). Federated learning in distributed medical databases: Meta-analysis of large-scale subcortical brain data. In 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019) (IEEE), pp. 270–274. https://doi.org/10.1109/ISBI.2019.8759317.

23. 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 (Springer), pp. 201–210. https://doi.org/10.1007/978-3-030-60548-3_20.

24. Li, X., Gu, Y., Dvornek, N., Staib, L.H., Ventola, P., and Duncan, J.S. (2020). Multi-site fmri analysis using privacy-preserving federated learning and domain adaptation: Abide results. Med. Image Anal. *65*, 101765. https://doi.org/10.1016/j.media.2020.101765.

25. Rootes-Murdy, K., Gazula, H., Verner, E., Kelly, R., DeRamus, T., Plis, S., Sarwate, A., Turner, J., and Calhoun, V. (2022). Federated analysis of neuroimaging data: a review of the field. Neuroinformatics *65*, 1–14. https://doi.org/10.1007/s12021-021-09550-7.

26. Stripelis, D., Saleem, H., Ghai, T., Dhinagar, N., Gupta, U., Anastasiou, C., Ver Steeg, G., Ravi, S., Naveed, M., Thompson, P.M., and Ambite, J.L. (2021). Secure neuroimaging analysis using federated learning with homomorphic encryption. In 17th International Symposium on Medical Information Processing and Analysis, *12088* (SPIE), pp. 351–359. https://doi.org/10.1117/12.2606256.

27. Mitrovska, A., Safari, P., Ritter, K., Shariati, B., and Fischer, J.K. (2024). Secure federated learning for alzheimer's disease detection. Front. Aging Neurosci. *16*, 1324032. https://doi.org/10.3389/fnagi.2024.1324032.

28. Fan, Z., Su, J., Gao, K., Hu, D., and Zeng, L.-L. (2021). A federated deep learning framework for 3d brain mri images. In 2021 International Joint Conference on Neural Networks (IJCNN) (IEEE), pp. 1–6. https://doi.org/10.1109/IJCNN52387.2021.9534376.

29. Stripelis, D., Gupta, U., Saleem, H., Dhinagar, N., Ghai, T., Anastasiou, C., Sanchez, R., Steeg, G.V., Ravi, S., Naveed, M., et al. (2024). Metisfl (University of Southern California. Zenodo). https://doi.org/10.5281/zenodo.11411754.

30. Stripelis, D., Anastasiou, C., Toral, P., Asghar, A., and Ambite, J.L. (2023). Metisfl: An embarrassingly parallelized controller for scalable & efficient federated learning workflows. In Proceedings of the 4th International Workshop on Distributed Machine Learning (Association for Computing Machinery), pp. 11–19. https://doi.org/10.1145/3630048.3630186.

31. Cheon, J.H., Kim, A., Kim, M., and Song, Y. (2017). Homomorphic encryption for arithmetic of approximate numbers. In Advances in Cryptology – ASIACRYPT 2017, T. Takagi and T. Peyrin, eds. (Springer), pp. 409–437. https://doi.org/10.1007/978-3-319-70694-8_15.

32. Geiping, J., Bauermeister, H., Dröge, H., and Moeller, M. (2020). Inverting gradients - how easy is it to break privacy in federated learning? In Advances in Neural Information Processing Systems, *33*, H. Larochelle, M. Ranzato, R. Hadsell, M. Balcan, and H. Lin, eds. *(Curran Associates, Inc.), pp. 16937–16947. https://proceedings.neurips.cc/paper_files/paper/2020/file/c4ede56bbd98819ae6112b20ac6bf145-Paper.pdf*.

33. Zhu, L., Liu, Z., and Han, S. (2019). Deep leakage from gradients. In Advances in Neural Information Processing Systems, *32*, H. Wallach, H. Larochelle, A. Beygelzimer, F. d' Alché-Buc, E. Fox, and R. Garnett, eds. *(Curran Associates, Inc). https://proceedings.neurips.cc/paper_files/paper/2019/file/60a6c4002cc7b29142def8871531281a-Paper.pdf*.

34. Shokri, R., Stronati, M., Song, C., and Shmatikov, V. (2017). Membership Inference Attacks Against Machine Learning Models. In 2017 IEEE Symposium on Security and Privacy (SP) (IEEE), pp. 3–18. https://doi.org/10.1109/SP.2017.41.

35. Nasr, M., Shokri, R., and Houmansadr, A. (2019). Comprehensive Privacy Analysis of Deep Learning: Passive and Active White-box Inference Attacks against Centralized and Federated Learning. In 2019 IEEE Symposium on Security and Privacy (SP) (IEEE), pp. 739–753. https://doi.org/10.1109/SP.2019.00065.

36. Miller, K.L., Alfaro-Almagro, F., Bangerter, N.K., Thomas, D.L., Yacoub, E., Xu, J., Bartsch, A.J., Jbabdi, S., Sotiropoulos, S.N., Andersson, J.L., et al. (2016). Multimodal population brain imaging in the uk biobank prospective epidemiological study. Nat. Neurosci. *19*, 1523–1536. https://doi.org/10.1038/nn.4393.

37. Mueller, S.G., Weiner, M.W., Thal, L.J., Petersen, R.C., Jack, C., Jagust, W., Trojanowski, J.Q., Toga, A.W., and Beckett, L. (2005). The alzheimer's disease neuroimaging initiative. Neuroimaging Clinics *15*, 869–877. https://doi.org/10.1016/j.nic.2005.09.008.

38. LaMontagne, P.J., Benzinger, T.L., Morris, J.C., Keefe, S., Hornbeck, R., Xiong, C., Grant, E., Hassenstab, J., Moulder, K., Vlassenko, A.G., et al. (2019). Oasis-3: Longitudinal neuroimaging, clinical, and cognitive dataset for normal aging and alzheimer disease. Preprint at medRxiv. https://doi.org/10.1101/2019.12.13.19014902.

39. Fowler, C., Rainey-Smith, S.R., Bird, S., Bomke, J., Bourgeat, P., Brown, B.M., Burnham, S.C., Bush, A.I., Chadunow, C., Collins, S., et al. (2021). Fifteen years of the australian imaging, biomarkers and lifestyle (aibl) study: progress and observations from 2,359 older adults spanning the spectrum from cognitive normality to alzheimer's disease. J. Alzheimers Dis. Rep. *5*, 443–468. https://doi.org/10.3233/ADR-210005.

40. Bonawitz, K., Ivanov, V., Kreuter, B., Marcedone, A., McMahan, H.B., Patel, S., Ramage, D., Segal, A., and Seth, K. (2017). Practical secure aggregation for privacy-preserving machine learning. In proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security (Association for Computing Machinery), pp. 1175–1191. https://doi.org/10.1145/3133956.3133982.

41. Joye, M., and Libert, B. (2013). A scalable scheme for privacy-preserving aggregation of time-series data. In Financial Cryptography and Data Security: 17th International Conference, FC 2013, Okinawa, Japan, April 1-5, 2013, Revised Selected Papers 17 (Springer), pp. 111–125. https://doi.org/10.1007/978-3-642-39884-1_10.

42. Hatamizadeh, A., Yin, H., Molchanov, P., Myronenko, A., Li, W., Dogra, P., Feng, A., Flores, M.G., Kautz, J., Xu, D., and Roth, H.R. (2023). Do gradient inversion attacks make federated learning unsafe? IEEE Trans. Med. Imag. *42*, 2044–2056. https://doi.org/10.1109/TMI.2023.3239391.

43. Huang, Y., Gupta, S., Song, Z., Li, K., and Arora, S. (2021). Evaluating gradient inversion attacks and defenses in federated learning. In Advances in Neural Information Processing Systems, *34*, M. Ranzato, A. Beygelzimer, Y. Dauphin, P. Liang, and J.W. Vaughan, eds. *(Curran Associates, Inc.), pp. 7232–7241. https://proceedings.neurips.cc/paper_files/paper/2021/file/3b3fff6463464959dcd1b68d0320f781-Paper.pdf*.

44. Gupta, U., Stripelis, D., Lam, P.K., Thompson, P., Ambite, J.L., and Steeg, G.V. (2021). Membership inference attacks on deep regression models for neuroimaging. In Proceedings of the Fourth Conference on Medical Imaging with Deep Learning vol. 143 of *Proceedings of Machine Learning Research*, M. Heinrich, Q. Dou, M. de Bruijne, J. Lellmann, A. Schläfer, and F. Ernst, eds. (PMLR), pp. 228–251. https://proceedings.mlr.press/v143/gupta21a.html.

45. Yeom, S., Giacomelli, I., Fredrikson, M., and Jha, S. (2018). Privacy risk in machine learning: Analyzing the connection to overfitting. In 2018 IEEE 31st computer security foundations symposium (CSF) (IEEE), pp. 268–282. https://doi.org/10.1109/CSF.2018.00027.

46. Truex, S., Liu, L., Gursoy, M.E., Yu, L., and Wei, W. (2018). Towards Demystifying Membership Inference Attacks. Preprint at arXiv. https://doi.org/10.48550/arXiv.1807.09173.

47. Salem, A., Zhang, Y., Humbert, M., Berrang, P., Fritz, M., and Backes, M. (2018). Ml-leaks: Model and data independent membership inference attacks and defenses on machine learning models. Preprint at arXiv. https://doi.org/10.48550/arXiv.1806.01246.

48. Jha, S.K., Jha, S., Ewetz, R., Raj, S., Velasquez, A., Pullum, L.L., and Swami, A. (2020). An Extension of Fano's Inequality for Characterizing Model Susceptibility to Membership Inference Attacks. Preprint at arXiv. https://doi.org/10.48550/arXiv.2009.08097.

49. Abadi, M., Chu, A., Goodfellow, I., McMahan, H.B., Mironov, I., Talwar, K., and Zhang, L. (2016). Deep Learning with Differential Privacy. In Proceedings of the 2016 ACM SIGSAC conference on computer and communications security (Association for Computing Machinery), pp. 308–318. https://doi.org/10.1145/2976749.2978318.

50. Foley, P., Sheller, M.J., Edwards, B., Pati, S., Riviera, W., Sharma, M., Moorthy, P.N., Wang, S.-h., Martin, J., Mirhaji, P., et al. (2022). Openfl: the open federated learning library. Phys. Med. Biol. *67*, 214001. https://doi.org/10.1088/1361-6560/ac97d9.

51. Roth, H.R., Cheng, Y., Wen, Y., Yang, I., Xu, Z., Hsieh, Y.-T., Kersten, K., Harouni, A., Zhao, C., Lu, K., et al. (2022). Nvidia flare: Federated learning from simulation to real-world. Preprint at arXiv. https://doi.org/10.48550/arXiv.2210.13291.

52. Beutel, D.J., Topal, T., Mathur, A., Qiu, X., Fernandez-Marques, J., Gao, Y., Sani, L., Li, K.H., Parcollet, T., de Gusmão, P.P.B., et al. (2022). Flower: A friendly federated learning framework. Preprint at arXiv. https://doi.org/10.48550/arXiv.2007.14390.

53. He, C., Li, S., So, J., Zeng, X., Zhang, M., Wang, H., Wang, X., Vepakomma, P., Singh, A., Qiu, H., et al. (2020). Fedml: A research library and benchmark for federated machine learning. Preprint at arXiv. https://doi.org/10.48550/arXiv.2007.13518.

54. Ludwig, H., Baracaldo, N., Thomas, G., Zhou, Y., Anwar, A., Rajamoni, S., Ong, Y., Radhakrishnan, J., Verma, A., Sinn, M., et al. (2020). Ibm federated learning: an enterprise framework white paper v0. 1. Preprint at arXiv. https://doi.org/10.48550/arXiv.2007.10987.

55. Kaissis, G., Ziller, A., Passerat-Palmbach, J., Ryffel, T., Usynin, D., Trask, A., Lima, I., Mancuso, J., Jungmann, F., Steinborn, M.-M., et al. (2021). End-to-end privacy preserving deep learning on multi-institutional medical imaging. Nat. Mach. Intell. *3*, 473–484. https://doi.org/10.1038/s42256-021-00337-8.

56. Galtier, M.N., and Marini, C. (2019). Substra: a framework for privacy-preserving, traceable and collaborative machine learning. Preprint at arXiv. https://doi.org/10.48550/arXiv.1910.11567.

57. Cremonesi, F., Vesin, M., Cansiz, S., Bouillard, Y., Balelli, I., Innocenti, L., Silva, S., Ayed, S.-S., Taiello, R., Kameni, L., et al. (2023). Fed-biomed: Open, transparent and trusted federated learning for real-world healthcare applications. Preprint at arXiv. https://doi.org/10.48550/arXiv.2304.12012.

58. Liu, Y., Fan, T., Chen, T., Xu, Q., and Yang, Q. (2021). Fate: An industrial grade platform for collaborative learning with data protection. J. Mach. Learn. Res. *22*, 1–6. http://jmlr.org/papers/v22/20-815.html.

59. Lai, F., Dai, Y., Zhu, X., Madhyastha, H.V., and Chowdhury, M. (2021). Fedscale: Benchmarking model and system performance of federated learning. Proceedings of the First Workshop on Systems Challenges in Reliable and Secure Federated Learning (PMLR), pp. 11814-11827. https://doi.org/10.1145/3477114.3488760.

60. 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. Front. Neurosci. *10*, 365. https://doi.org/10.3389/fnins.2016.00365.

61. Li, Q., Wen, Z., Wu, Z., Hu, S., Wang, N., Li, Y., Liu, X., and He, B. (2023). A survey on federated learning systems: vision, hype and reality for data privacy and protection. IEEE Trans. Knowl. Data Eng. *35*, 3347–3366. https://doi.org/10.1109/TKDE.2021.3124599.

62. Kairouz, P., McMahan, H.B., Avent, B., Bellet, A., Bennis, M., Bhagoji, A.N., Bonawitz, K., Charles, Z., Cormode, G., Cummings, R., et al. (2021). Advances and open problems in federated learning. FNT. in Machine Learning *14*, 1–210. https://doi.org/10.1561/2200000083.

63. Liu, X., Shi, T., Xie, C., Li, Q., Hu, K., Kim, H., Xu, X., Li, B., and Song, D. (2022). Unifed: A benchmark for federated learning frameworks. Preprint at arXiv. https://doi.org/10.48550/arXiv.2207.10308.

64. Liu, Y., Kang, Y., Zou, T., Pu, Y., He, Y., Ye, X., Ouyang, Y., Zhang, Y.-Q., and Yang, Q. (2024). Vertical federated learning: Concepts, advances, and challenges. IEEE Trans. Knowl. Data Eng. *36*, 3615–3634. https://doi.org/10.1109/TKDE.2024.3352628.

65. Naseri, M., Hayes, J., and De Cristofaro, E. (2020). Local and central differential privacy for robustness and privacy in federated learning. Preprint at arXiv. https://doi.org/10.48550/arXiv.2009.03561.

66. Jin, W., Yao, Y., Han, S., Joe-Wong, C., Ravi, S., Avestimehr, S., and He, C. (2023). Fedml-he: An efficient homomorphic-encryption-based privacy-preserving federated learning system. Preprint at arXiv. https://doi.org/10.48550/arXiv.2303.10837.

67. So, J., He, C., Yang, C.-S., Li, S., Yu, Q., E. Ali, R., Guler, B., and Avestimehr, S. (2022). Lightsecagg: a lightweight and versatile design for secure aggregation in federated learning. In Proceedings of Machine Learning and Systems, *4*, D. Marculescu, Y. Chi, and C. Wu, eds.*, pp. 694–720. https://proceedings.mlsys.org/paper_files/paper/2022/file/6c44dc73014d66ba49b28d483a8f8b0d-Paper.pdf*.

68. Li, K.H., de Gusmão, P.P.B., Beutel, D.J., and Lane, N.D. (2021). Secure aggregation for federated learning in flower. In Proceedings of the 2nd ACM International Workshop on Distributed Machine Learning (Association for Computing Machinery), pp. 8–14. https://doi.org/10.1145/3488659.3493776.

69. Damgård, I., Pastro, V., Smart, N., and Zakarias, S. (2012). Multiparty computation from somewhat homomorphic encryption. In Annual Cryptology Conference (Springer), pp. 643–662. https://doi.org/10.1007/978-3-642-32009-5_38.

70. Zhang, C., Li, S., Xia, J., Wang, W., Yan, F., and Liu, Y. (2020). BatchCrypt: Efficient homomorphic encryption for Cross-Silo federated learning. In 2020 USENIX Annual Technical Conference (USENIX ATC 20) (USENIX Association), pp. 493–506. https://www.usenix.org/conference/atc20/presentation/zhang-chengliang.

71. Sabt, M., Achemlal, M., and Bouabdallah, A. (2015). Trusted execution environment: What it is, and what it is not. In 2015 IEEE Trustcom/BigDataSE/ISPA, *1* (IEEE), pp. 57–64. https://doi.org/10.1109/Trustcom.2015.357.

72. Stripelis, D., Thompson, P.M., and Ambite, J.L. (2022). Semi-synchronous federated learning for energy-efficient training and accelerated convergence in cross-silo settings. ACM Trans. Intell. Syst. Technol. *13*, 1–29. https://doi.org/10.1145/3524885.

73. Dagum, L., and Menon, R. (1998). Openmp: an industry standard api for shared-memory programming. IEEE Comput. Sci. Eng. *5*, 46–55. https://doi.org/10.1109/99.660313.

74. Patterson, C. (2018). World Alzheimer Report 2018 (Alzheimer's Disease International).

75. Dhinagar, N., Thomopoulos, S.I., Owens-Walton, C., Stripelis, D., Ambite, J.L., Ver Steeg, G., Weintraub, D., Cook, P., McMillan, C., and Thompson, P.M. (2021). 3D convolutional neural networks for classification of Alzheimer's and Parkinson's disease with T1-weighted brain MRI. In 17th International Symposium on Medical Information Processing and Analysis, *12088*, L. Rittner, M.D.,E.R.C., N. Lepore, J. Brieva, and M.G. Linguraru, eds. *(International Society for Optics and Photonics), pp. 277–286. https://doi.org/10.1117/12.2606297*.

76. AbdulAzeem, Y., Bahgat, W.M., and Badawy, M. (2021). A cnn based framework for classification of alzheimer's disease. Neural Comput. Appl. *33*, 10415–10428. https://doi.org/10.1007/s00521-021-05799-w.

77. Fu'adah, Y., Wijayanto, I., Pratiwi, N., Taliningsih, F., Rizal, S., and Pramudito, M. (2021). Automated classification of alzheimer's disease based on mri image processing using convolutional neural network (cnn) with alexnet architecture. In Journal of Physics: Conference Series, *1844* (IOP Publishing), pp. 012020. https://doi.org/10.1088/1742-6596/1844/1/012020.

78. Franke, K., and Gaser, C. (2019). Ten years of brainage as a neuroimaging biomarker of brain aging: what insights have we gained? Front. Neurol. *10*, 789. https://doi.org/10.3389/fneur.2019.00789.

79. Wood, D.A., Kafiabadi, S., Al Busaidi, A., Guilhem, E., Montvila, A., Lynch, J., Townend, M., Agarwal, S., Mazumder, A., Barker, G.J., et al. (2022). Accurate brain-age models for routine clinical mri examinations. Neuroimage *249*, 118871. https://doi.org/10.1016/j.neuroimage.2022.118871.

80. Cole, J.H., Leech, R., Sharp, D.J., and Initiative, A.D.N. (2015). Prediction of brain age suggests accelerated atrophy after traumatic brain injury. Ann. Neurol. *77*, 571–581. https://doi.org/10.1002/ana.24367.

81. Koutsouleris, N., Davatzikos, C., Borgwardt, S., Gaser, C., Bottlender, R., Frodl, T., Falkai, P., Riecher-Rössler, A., Möller, H.-J., Reiser, M., et al. (2014). Accelerated brain aging in schizophrenia and beyond: a neuroanatomical marker of psychiatric disorders. Schizophr. Bull. *40*, 1140–1153. https://doi.org/10.1093/schbul/sbt142.

82. Kuchinad, A., Schweinhardt, P., Seminowicz, D.A., Wood, P.B., Chizh, B.A., and Bushnell, M.C. (2007). Accelerated brain gray matter loss in fibromyalgia patients: premature aging of the brain? J. Neurosci. *27*, 4004–4007. https://doi.org/10.1523/JNEUROSCI.0098-07.2007.

83. Lam, P.K., Santhalingam, V., Suresh, P., Baboota, R., Zhu, A.H., Thomopoulos, S.I., Jahanshad, N., and Thompson, P.M. (2020). Accurate brain age prediction using recurrent slice-based networks. In 16th International Symposium on Medical Information Processing and Analysis, *11583* (International Society for Optics and Photonics), pp. 11–20. https://doi.org/10.1117/12.2579630.

84. Jónsson, B.A., Bjornsdottir, G., Thorgeirsson, T., Ellingsen, L.M., Walters, G.B., Gudbjartsson, D., Stefansson, H., Stefansson, K., and Ulfarsson, M. (2019). Brain age prediction using deep learning uncovers associated sequence variants. Nat. Commun. *10*, 1–10. https://doi.org/10.1038/s41467-019-13163-9.

85. Dinsdale, N.K., Bluemke, E., Smith, S.M., Arya, Z., Vidaurre, D., Jenkinson, M., and Namburete, A.I. (2021). Learning patterns of the ageing brain in mri using deep convolutional networks. Neuroimage *224*, 117401. https://doi.org/10.1016/j.neuroimage.2020.117401.

86. Cole, J.H., Poudel, R.P., Tsagkrasoulis, D., Caan, M.W., Steves, C., Spector, T.D., and Montana, G. (2017). Predicting brain age with deep learning from raw imaging data results in a reliable and heritable biomarker. Neuroimage *163*, 115–124. https://doi.org/10.1016/j.neuroimage.2017.07.059.

87. Peng, H., Gong, W., Beckmann, C.F., Vedaldi, A., and Smith, S.M. (2021). Accurate brain age prediction with lightweight deep neural networks. Med. Image Anal. *68*, 101871. https://doi.org/10.1016/j.media.2020.101871.

88. Gupta, U., Lam, P.K., Ver Steeg, G., and Thompson, P.M. (2021). Improved brain age estimation with slice-based set networks. In 2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI) (IEEE), pp. 840–844. https://doi.org/10.1109/ISBI48211.2021.9434081.

89. Stripelis, D., Ambite, J.L., Lam, P., and Thompson, P. (2021). Scaling Neuroscience Research Using Federated Learning. In 2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI) (Ieee), pp. 1191–1195. https://doi.org/10.1109/ISBI48211.2021.9433925.

90. Albrecht, M., Chase, M., Chen, H., Ding, J., Goldwasser, S., Gorbunov, S., Halevi, S., Hoffstein, J., Laine, K., Lauter, K., et al. (2019). Homomorphic Encryption Standard (Cryptology ePrint Archive). https://eprint. iacr.org/2019/939.

91. Sako, K. (2011). Public Key Cryptography (Springer US), pp. 996–997. https://doi.org/10.1007/0-387-23483-7_331.

92. Truex, S., Baracaldo, N., Anwar, A., Steinke, T., Ludwig, H., Zhang, R., and Zhou, Y. (2019). A hybrid approach to privacy-preserving federated learning. In Proceedings of the 12th ACM Workshop on Artificial Intelligence and Security (AISec'19 New York, NY, USA: Association for Computing Machinery), pp. 1–11. https://doi.org/10.1145/3338501.3357370.

93. Ma, J., Naas, S.-A., Sigg, S., and Lyu, X. (2022). Privacy-preserving federated learning based on multi-key homomorphic encryption. Int. J. Intell. Syst. *37*, 5880–5901. https://doi.org/10.1002/int.22818.

94. Wei, K., Li, J., Ding, M., Ma, C., Yang, H.H., Farokhi, F., Jin, S., Quek, T.Q.S., and Vincent Poor, H. (2020). Federated learning with differential privacy: Algorithms and performance analysis. IEEE Trans. Inf. Forensics Secur. *15*, 3454–3469. https://doi.org/10.1109/TIFS.2020.2988575.

95. Noble, M., Bellet, A., and Dieuleveut, A. (2022). Differentially private federated learning on heterogeneous data. In Proceedings of The 25th International Conference on Artificial Intelligence and Statistics vol. 151 of *Proceedings of Machine Learning Research*, G. Camps-Valls, F.J.R. Ruiz, and I. Valera, eds. (PMLR), pp. 10110–10145. https://proceedings.mlr.press/v151/noble22a.html.

96. Zhao, Y., Zhao, J., Yang, M., Wang, T., Wang, N., Lyu, L., Niyato, D., and Lam, K.-Y. (2021). Local differential privacy-based federated learning for internet of things. IEEE Internet Things J. *8*, 8836–8853. https://doi.org/10.1109/JIOT.2020.3037194.

97. Veitch, D.P., Weiner, M.W., Aisen, P.S., Beckett, L.A., Cairns, N.J., Green, R.C., Harvey, D., Jack Jr, C.R., Jagust, W., Morris, J.C., et al. (2019). Understanding disease progression and improving alzheimer's disease clinical trials: Recent highlights from the alzheimer's disease neuroimaging initiative. Alzheimers Dement. *15*, 106–152. https://doi.org/10.1016/j.jalz.2018.08.005.

98. Stripelis, D., and Ambite, J.L. (2020). Accelerating federated learning in heterogeneous data and computational environments. Preprint at arXiv. https://doi.org/10.48550/arXiv.2008.11281.

99. Stripelis, D., Abram, M., and Ambite, J.L. (2022). Performance weighting for robust federated learning against corrupted sources. Preprint at arXiv. https://doi.org/10.48550/arXiv.2205.01184.

100. Wang, J., Charles, Z., Xu, Z., Joshi, G., McMahan, H.B., Al-Shedivat, M., Andrew, G., Avestimehr, S., Daly, K., Data, D., et al. (2021). A field guide to federated optimization. Preprint at arXiv. https://doi.org/10.48550/arXiv.2107.06917.

101. Reddi, S.J., Charles, Z., Zaheer, M., Garrett, Z., Rush, K., Konečný, J., Kumar, S., and McMahan, H.B. (2021). Adaptive federated optimization. In 9th International Conference on Learning Representations, ICLR 2021, Virtual Event, Austria, May 3-7, 2021 (OpenReview.net) https://openreview.net/forum?id=LkFG3lB13U5.

102. Hsu, T.-M.H., Qi, H., and Brown, M. (2019). Measuring the effects of non-identical data distribution for federated visual classification. Preprint at arXiv. https://doi.org/10.48550/arXiv.1909.06335.

103. Moyer, D., Ver Steeg, G., Tax, C.M., and Thompson, P.M. (2020). Scanner invariant representations for diffusion mri harmonization. Magn. Reson. Med. *84*, 2174–2189. https://doi.org/10.1002/mrm.28243.

104. Komandur, D., Gupta, U., Chattopadhyay, T., Dhinagar, N.J., Thomopoulos, S.I., Chen, J.-C., Beavers, D., Ver Steeg, G., and Thompson, P.M. (2023). Unsupervised harmonization of brain mri using 3d cyclegans and its effect on brain age prediction. In 2023 19th International Symposium on Medical Information Processing and Analysis (SIPAIM) (IEEE), pp. 1–5. https://doi.org/10.1101/2022.11.15.516349.

105. Dwork, C., and Roth, A. (2013). The Algorithmic Foundations of Differential Privacy. FNT. in Theoretical Computer Science *9*, 211–407. https://doi.org/10.1561/0400000042.

106. Jayaraman, B., and Evans, D. (2019). Evaluating differentially private machine learning in practice. In 28th USENIX Security Symposium (USENIX Security 19) (USENIX Association), pp. 1895–1912. https://www.usenix.org/conference/usenixsecurity19/presentation/jayaraman.