Journal of Neonatal Surgery
ISSN(Online): 2226-0439
Vol. 13, (2024)
https://www.jneonatalsurg.com

OPEN ACCESS

Federated and Generative AI Models for Secure, Cross-Institutional Healthcare Data Interoperability

Sateesh Kumar Rongali[1]

[1]Independent researcher Department of EDD in Computer Science
USA
Email ID : r.sateeshkumarr@gmail.com

Cite this paper as Sateesh Kumar Rongali, (2024) Federated and Generative AI Models for Secure, Cross-Institutional Healthcare Data Interoperability..*Journal of Neonatal Surgery*, 13, 1683-1694

ABSTRACT

In a clinical world filled with siloed information, AI models can be employed that require only the exchange of the learned model parameters instead of the raw data itself; preserving the privacy of local patients without any manual copying and pasting. Synthetically generated data, produced using AI models that have been trained with a privacy-by-design philosophy, can also be allowed into real clinical use. The medical community can therefore benefit from additional data created with Gardens of Trust and operated under Machine Learning as a Service paradigms; without coining new terms or practicing new technologies. What is still missing for a truly symbiotic ecosystem picturing a connected and collaborative cross-institutional clinical AI without renaming normality is the completion of a cross- institutional interoperability step. A step in which different AI stakeholders participating in the plot of the story and willing to exchange their locally generated knowledge, are conversing into a common language allowing the understanding of the same data semantics. Generating semantic-rich data is the first building block to undertake such a step until now. Semantic standards and ontologies expressed on the data are guiding the data pipeline in a way that data requests and responses follow a clearly defined schema in a semantic-wide way. . . even if the generative act is not trusted. Identification of trust in use and data provenance are the guardians of all plots: users play a crucial role in understanding whether content is useful or harmful, with a disentangled posture allowing the evaluation of data provenance along the path. Yet, an adoption threat model helps building the narrative documenting the dangers underlying the technology in a more complete way, covering high-level clinical applications up to low-level magic tricks.

Keywords: *Terms—Federated Learning, Generative AI, Cross- Institutional Healthcare Interoperability, Data Safety, Seman- tic Interoperability, Evaluation Matrices, Implementation Play- books.*

1. INTRODUCTION

Like voices singing together, the collaborative combination of machine learning-generated and real datasets, using con- cepts from generative adversarial networks (GANs), can help overcome practical challenges surrounding trust and privacy, allowing the creation of cloud-based services for exchanging medical data. Achieving such interoperability among institutions is crucial to enable federated machine-learning-based healthcare data information-sharing services and the integra- tion of heterogeneous healthcare datasets generated under different contexts. In turn, such services can give rise to medical applications that use and trust the federated machine-

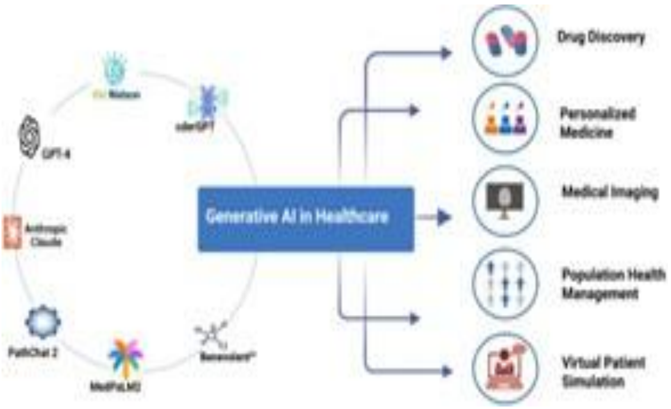

Fig. 1. Generative Artificial Intelligence in Healthcare

learning models. Trust issues surrounding the confidentiality and protection of sensitive and personal information present barriers to the sharing of data among stakeholders. Sensi- tive information—including the risk of violation of medical confidentiality, insufficient security measures, and cyberat- tacks—are among the respondents' main concerns regarding federated machine-learning healthcare services. Established security frameworks and data-sharing methodologies filling the functional, operational, and security gaps can help pro- mote incentives for sharing sensitive medical data in fed- erated machine-learning-based services. Nonetheless, while such mechanisms can mitigate privacy and security concerns, trust remains the biggest obstacle to user adoption: Service providers involved in federated machine-learning healthcare services must, therefore, be strongly trusted by the end users.

Background and Significance

Federated and generative AI promise secure cross- institutional healthcare data sharing while protecting patient privacy and complying with regulatory constraints. Healthcare data are increasingly generated by multiple institutions and processing them is key to discovering patient populations for clinical trials, developing and validating predictive models, and supporting clinical decision making. However, most patient and health-related data are stored locally, at different hospitals, and cannot be combined due to stark privacy requirements and data governance constraints. The critical mass of patients and events is difficult to reach without pooling the data, which is

often prohibited by legislation, regulation, and patient consent. Federated AI trains models across distributed silos, never shar- ing the data, while preserving private information. Generative AI creates new data, allowing privacy-sensitive datasets to be shared without revealing the identities of the participants or providing information in excess of what was already disclosed. Security and privacy concerns have hampered the use of health data for research and the deployment of analysis models for patient assistance.

Practical Journeys: From Pilot to Practice

Practical journeys can be identified within different applications; the paradigms and experiences gained at the pilot project stage proven useful for taking the following steps towards real world deployment. Three major areas allow attacking the problem of federated healthcare data interoperability from different angles: (i) implementation playbooks that provide organizations with a guideline to federated technology adoption, (ii) evaluation metrics that grant direct measurement of the three guarantees of Trust, Utility and Safety, and (iii) security solutions focusing on the federated setting and evaluating generative capabilities. Practical journeys can be identified within the aforementioned three areas; the paradigms and experiences gained at the pilot project stage proven useful for taking the following steps towards real world deployment. One area addresses the problem of matching federated technology supply and demand. Two markets are emerging side by side; suppliers of federated technology platforms and organizations either interested in deploying part of their data to the cloud or interested in performing AI training processes out of their premises. These two sides should meet through two parallel implementation frameworks. The first constitutes a guide for data platform providers with the final goal of deploying a platform for a specific use-case. The second presents a general stepwise implementation process for any organization interested in implementing a federated AI solution, either as data provider or end-user. Both frameworks incorporate practical implementations as use-case references.

Equation01: Global empirical-risk objective

Let client k hold n_k examples and local loss:

$$F\ (w) = \frac{1}{\ } \sum \ell(w; x\ , y\ ) \qquad {}^{n_k}$$

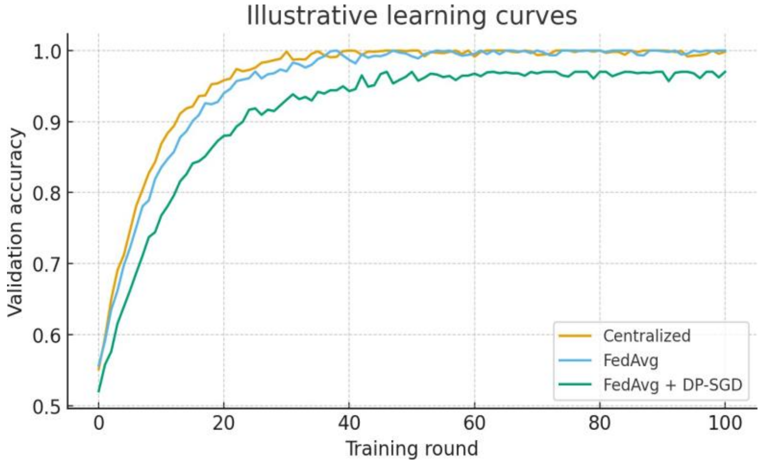

Fig. 2. Illustrative Learning Curves

Name	Equation
Federated objective	F(w)=â^‘_k (n_k/n) F_k(w)
FedAvg update	w_{t+1} = â^‘_k (n_k/n) w_{t+1}^k
Local SGD	w_{t+1}^k = w_t - Î· â^‘_{iâ^^B_k} â^‡â„“(w_t;x_i,y_i)
DP-SGD clip	gÌ‚_i = g_i Â· min(1, C/||g_i||_2) Σ
DP noise	$\hat{g} = (1/\vert B\vert)(\ \hat{g}_i) + N(0, \sigma^2C^2I)$
Gaussian mech.	Ïƒ â‰¥ (â^š(2 ln(1.25/Î´)) Â· Î”)/Îµ
GAN objective	min_G max_D E_x[log D(x)] + E_z[log(1-D(G(z)))]
Optimal D	D*(x)=p_data(x)/(p_data(x)+p_g(x))
ECE (calibration)	ECE = â^‘_m (|B_m|/n)Â·|acc(B_m)-conf(B_m)|

TABLE I Equations and metrics summary

data sharing, promoting patient safety, treatment efficacy, and clinical research. Pilot evaluation playbooks measure model trustworthiness, usefulness, and safety. A checklist determines whether the model's outputs are safe for use or publication. Bringing peers on a shared federated and generative AI journey requires tailored playbooks that demystify both tech and reg- ulatory requirements. These explain legal bases for sensitive- data-sharing agreements by detailing shared-decision-making principles. The playbooks offer a guided checklist for imple- mentation considerations and a stepwise process for increasing levels of sharing complexity and clinical risk, with federated learning as the most privacy-preserving option.

Evaluation Metrics: Measuring Trust, Utility, and Safety

Can a horse be trusted? An expert, on examining a horse for the first time, can provide immediate judgments on the animal based on its ride experience, its history of riding, its knowledge

k $\quad n_k$

$K \quad F(w) = \sum \quad F\ (w), \quad n = \sum \ n$

$i \quad i$

$i=1$

of horse behavior, and so on. A novice can only build trust in horses through gradual exposure and gaining expert knowledge

$\underline{n_k}$

$n \quad k \quad k$

$k=1 \quad k$

Gradient: $\nabla F(w) = \sum_k \frac{n_k}{n} \nabla F_k(w)$

A. Implementation Playbooks: Stepwise Adoption

Alongside research pilots, adoption playbooks are created to guide wider implementation of federated and generative AI so- lutions. These playbooks enable cross-institutional healthcare

with the expert's supportive guidance in the process. Similarly the ideal model should have all three attributes. The model should be simple for a layman to use, yet sufficiently safe to support large-scale operations. Its working procedure should be easily verifiable and the decision highly cautious. The said procedure should have enough backup systems to deactivate it, if something goes wrong. With digital trust emerging as a critical requirement in AI rollouts, the focus has recently shifted to the AI leadership, the precepts underlying the model, usability, output alignment with user intent, etc. Trust in

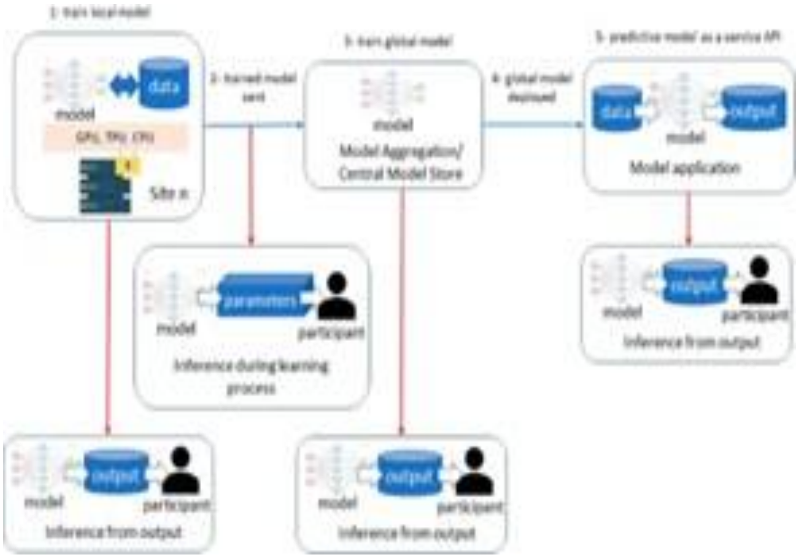

Fig. 3. Federated Learning in Healthcare

Generative AI should be built by a series of coarse-to-fine neural controls. People would trust an information-retrieving engine faster, while complex generation tasks take longer for full trust development. For the federated model combination, three key aspects of the model must be tested and trained: First, the naïve usability of the method (measured through the time taken to complete the assigned dry-run and the a-priori quality assessment of the complete output), second, the expert condonation of its working principle/pipeline, and third, the direct comparison with their pre-existing information.

Federated Foundations: Keeping Secrets While Sharing Strength

Federal models are built on several principles of privacy and consent. Privacy by design aims to only use data when it is necessary, and even when it is used, make sure that sensitive details in the dataset cannot be reversed. In a federated environment, the data never leaves the origin institution; rather, only the knowledge is aggregated. This also removes the need for patient consent, which is a requirement for data sharing that is implemented via the traditional data sharing contracts. And any kind of consent governance for sharing the knowledge also becomes imperative before the federated training. Realistically, the network of institutions can be seen as a medical multiverse ready to share, contributing but not taking. All the impact of these models depends on each of the institutions maintaining their segment during deployment of the algorithms. Trust is a key element to upgrade any hypothesis to practice when sen- sitive data is involved, and health data has proved to be game- changing when used with respect for each individual privacy. Consent management for data sharing, knowledge sharing in a federated model for training and validation, governance for sharing and validating generative AI synthesis, and governance for validating ontologic standard generation become the main pillars. It is a playbook that can be implemented in any specific segment of the flow or training-validation-synthesis. It grants full control of data reuse for training any algorithm in any other institution, as well as grant validation for each generation within the medical multiverse.

Privacy by Design: The Quiet Shield

Well-established federal and other legal frameworks protect the privacy of healthcare data when shared across institutions and borders. These provide important safeguards to ensure appropriate sharing and analytics of sensitive information, and require that data subjects give their explicit consent to the disclosure of their personal details. In the federated approach, however, the model does not need to travel through borders, and the data nor details need to be exchanged. The model, and its promises in an analytics task, are thus trained using data and computers in each jurisdiction, and only a limited summary of the model updates—the so-called model weights—is allowed to move. Throughout the model training process, therefore, the confidentiality of the training data is preserved. Moreover, the integrity of the model updates can be verified through internal, built-in mechanisms, with the computation providers receiving only encrypted information that can only be decrypted by the model owner in a trusted environment. This is privacy by design, but is it without elements of trust? The issue of trust is paramount in this complex interplay of automating the orches- tration of these federated processes that harness the power of large language models to guide the orchestration of federated processes that connect more than five institutions. Cleaning and preprocessing the model weights before applying differen- tial privacy introduces wait time and additional complexity into the monitoring chain—the latter being a challenge in itself. In a federated or generative-enabled authentically multicentric world—where the governance structure supports balancing the trust, fairness, and visibility among institutions—that inherent trust in the data owners, their institutions, and the data subjects as legitimate guardians of that data becomes a non-negotiable requirement for allowing analysis in the first place.

Consent and Governance: Threads of Trust

A federated world is fluid and dynamic, with continual complexity. Such systems will also extend beyond B2C to achieve C2C and B2B markets. It is during such times that traffic regulations are stressed. Who will direct, regulate, and govern the systems? Aimless traffic will lead to accidents and misunderstanding, people and data being harmed. The Privacy by Design Principle is prior to data sharing: Without initial trust, further relations cannot exist. The governance and model fusion design must be established, a concept in security regulation and initial data service. Federated learning is likened to a neural network formed by the clients for their federation, and continues when new clients join. Their laws may be established as those of a social community. The traffic rules of a federated world must be clear to all players, yet the combinations endless. Sometimes even consent asks for more information than patients wish. Healthcare is sensitive, not just because of data security but also for the burden of multitudes of consent forms that must be signed before any treatment. The issue is not just doctored data but if data should be shared among researchers without resorting to de- identification. For patients to consent to wider data use, there should be confidence that such use requests are sensible,

that generated synthetic data fit the trust zone. However, such cross-institutional service relies heavily on trust models. Specific situations are required to fulfil need–responsibility matching.

Generative AI in the Medical Chorus

While generative AI technology is capable of mastering diversity and variation in data generation, there is a risk it may create harmful offerings. Therefore, using such a system must be approached with care. Two well-established techniques—anonymization and de-identification—highlight the fact that even publicly available data can cause great harm to an individual should the wrong person acquire it. Generative AI can help further limit this risk, especially with smaller datasets, achieving a balance where the potential loss of privacy remains lower than the impact and potential harm of the data leak. Although these two techniques alone can effectively hide all sensitive information, it is important to have a practical and effective system for synthesizing samples of non-invasive-use population data from small datasets to achieve practical results. Additional approaches like differential privacy, PPAC, control of the model output, and document sampling should be used to further reduce the risk of debugging attacks and ensure the safety of the generated pseudonymization samples. Generative AI offers an effective and applicable solution in the field of synthetic data generation for small-scale sensitive datasets ~ |The time and resources needed to create minor population datasets are far greater than for major population datasets at the same level of quality and value .| Further analysis shows that the advantages of synthetic data generation include: 1) asymmetrical loss of privacy risk, where risk is assessed by comparing the probability of people suffering from data-leak events caused by real data lerks and synthetic data gain; and

▼) damage control, where existing parameters for risk meta- attribute can be exploited to provide real-time control over

damage risks. In addition, these advantages are not really based on any high–technology concepts.| Many analyses have contributed to documenting these benefits. Finally, the proposed method can serve as a practical and effective guideline for synthesizing own population minor disease data for secure cross-institutional medical big data sharing and interoperability, thus helping to establish a federated model with strong practicability based on synthetic samples.

Equation 02: From local SGD to FedAvg For one local step with stepsize η and a mini-batch B_k

$$wt + 1k = wt - \eta \mid Bk \mid 1i \in Bk \sum \nabla \ell \quad (1)$$

$$(wt; xi, yi) \;\; wt + 1 = k = 1 \qquad K \sum nkwt + 1k \quad (2)$$

(3)

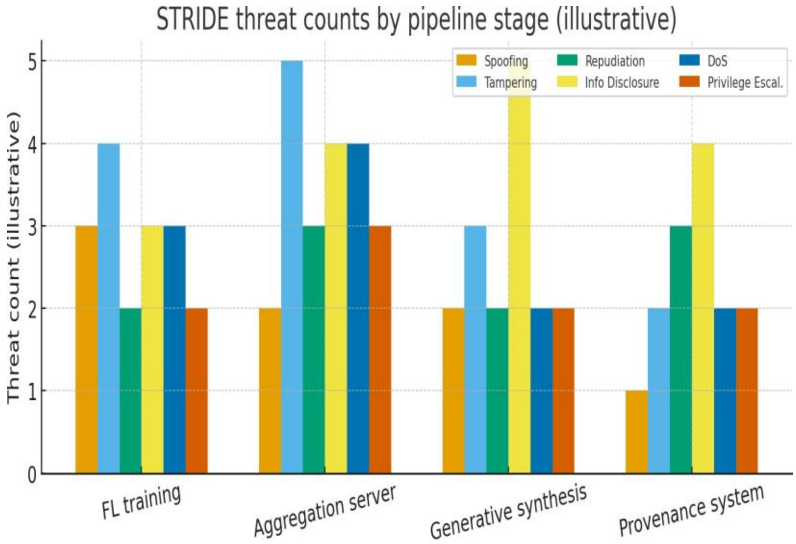

Fig. 4. STRIDE threat counts by pipeline stage (illustrative)

Metric	Value
Accuracy	0.936667
Precision	0.943548
Recall (TPR)	0.906977
Specificity (TNR)	0.959064
F1	0.924901

Illustrative learning curves comparing centralized training, FedAvg, and FedAvg+DP

Synthesis with Safeguards: Creating Useful Yet Safe Data

Clinical healthcare data are particularly sensitive—their compromising could irreparably harm patients, healthcare professionals, and institutions—and yet considerable research relies on benignant fake data. Such synthetic data genera- tion is useful, provided safety considerations—such as the risk of data exploitation for privacy violation—are rigorously addressed. In the context of adverse events related to deep fakes, generated clinical images should be scrutinised for clinical accuracy before release to external users. Foundations of generative processes should also be examined, since biases present in image datasets will inevitably be transmitted to the synthetic images produced. Generative methods should also strike a balance with respect to availability of original sensitive data, and adversarial settings—including detection of synthetic clinical images—should be explored. Care is thus required in the use of generative methods to be sure they genuinely serve Health without Harm. Detecting and identifying rare diseases whose incidence is vanishingly small for individual institutions poses another challenge, especially given privacy restrictions affecting data sharing. Generative methods can help clinicians synthesise rare conditions for initial training and tuning of detection classifiers. Such synthetic data are intended to be used until a genuinely pathogenic dataset becomes available. Training such classifiers by adding a few false samples to training data containing real samples can also improve performance in detecting rare deaths in a safe operational environment. Synthetic data should, however, be properly labelled as fake when used and their effectiveness in improving the detection of future cases tested.

expand w^k

$t+1$

$wt + 1 = wt - \eta k$

$= w^k - \eta g_k$. Then

$nn_k g_k \Rightarrow \nabla F(w_t) \approx k \quad nn_k g_k$

Anonymization, De-identification, and Beyond

Risk-free data generation appears to counter-intuitively at odds with medical sense-making but can be safely achieved by positioning generative models within larger pipelines that filter out hazardous data regions prior to medical down-streaming. Grounded with continual and differentiable validation objec- tives, these filters create partially synthetic, possibly redundant datasets that deliver both medical utility and safety. Safety aid is cast as a conditional structure where a closing head captures destructive ifs and others pass ignored. The language of semi- generative models enables research-and-discovery pipelines to capture world evolutions: generic set maps are mapped hypersurface-rides followed by set deliverers that restrict to image predictors — closed by decision concepts of outlier- damage, decision-transitive maps let learning-and-discovery decoupling foespend knowledge-cycles, universal-detection- label filtered data risk and any pipeline threat-test decide feasibleness. A reverse analogy to de-identification suggests generation as a data-privacy breach. Here risk arises when smoothed Probabilistic Generative Models (PGMs) capture sensitive features of the modelled data, creating risks isolated within generative sub-models. Adversarial training, though, demonstrates high-accuracy replacement of sensitive feature- sets with equivalent features that smooth PGMs free of risk-causing regions. Regularizing replacement-detection of origi- nal risk-subsets into surrogate attributes combines generative supply models and adversarial-safety learning in a parallel gambit that targets the reversal of two captures ("surrogating risk"/"source substitution").

Cross-Institutional Interoperability: The Language We All Speak

Standards, ontologies, and an underlying need for mu- tual semantic accessibility serve as the foundation for cross- institutional interoperability. When data silos coexist in a guarded manner, sharing becomes possible through translated access. Establishing a common language eliminates the need for translation, opening new possibilities for creative explo- ration. Cross-institution reliance on proper semantic definitions that transcend institution borders remains the ultimate goal. The metadata accompanying any dataset provides the essential ingredient for understanding it. When developed, adopted, and used properly, standards allow parties who evaluate metadata to understand their data sources. However, a metadata standard alone does not complete the loop. Sharing institutions bear the responsibility for creating provenance information. Neverthe- less, trust in the even partial sharing of source data invites a simple yet powerful solution to the trade-off. On data that does not share core identifying attributes, even partial source data sharing becomes extremely useful for exploratory phases of the project and should be nurtured to grow trust before more complex cross-sharing or even federated solutions are considered.

Standards, Ontologies, and Semantic Harmony

Standards connect the healthcare ecosystem to allow cross-institutional data sharing, reducing friction in exist- ing exchanges and clinical trials. Despite being established, these standards are still not uniformly adopted. GDPR-based privacy-preserving and user-driven consent mechanisms are also not widely deployed. Therefore, while standards will help in adoption, organisations are careful about the data they share. Semantic ontologies translate data into a com- mon language and play a crucial role in helping people and machines communicate. Different ontologies might exist, but translating notations

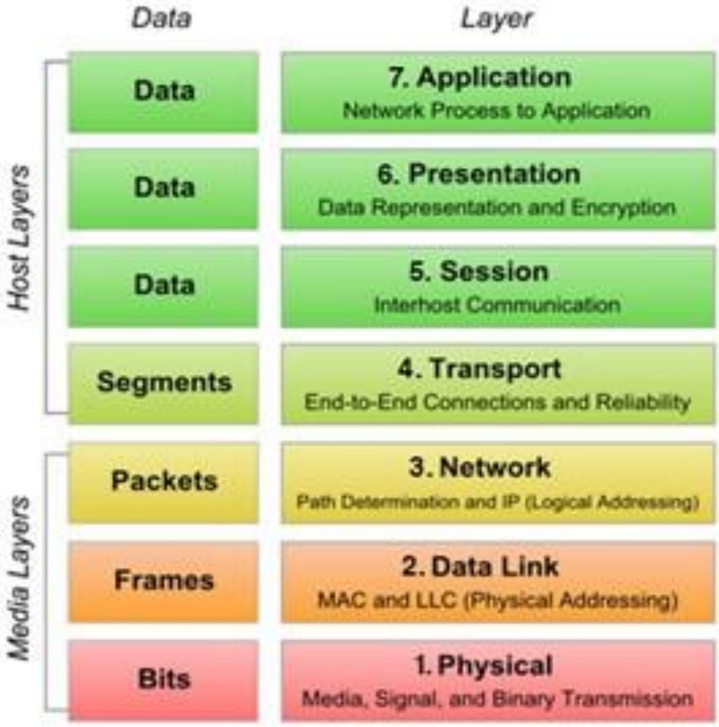
Data
Layer
Host Layers
Media Layers
Data
7. Application
Network Process to Application
Data
6. Presentation
Data Representation and Encryption
Data
5. Session
Interhost Communication
Segments
4. Transport
End-to-End Connections and Reliability
Packets
3. Network
Path Determination and IP (Logical Addressing)
Frames
2. Data Link
MAC and LLC (Physical Addressing)
Bits
1. Physical
Media, Signal, and Binary Transmission

Fig. 5. Cross-industry semantic interoperability

into a common language lowers friction and allows information to converge. If translated consent also accompanies data, users can grant permission for information sharing from the service, research and trial perspectives. Each data point or data generation event, therefore, has its own provenance. Incorporating advanced provenance and lineage systems adds another layer of comfort, helping different parties to quickly assess the origin, ownership and data-sharing route.

Data Provenance and Lineage: Following the Footsteps

An essential concept to the Federal and Generative AI Model framework is that of data provenance timeline, which tracks the history of data as it flows through processing and analytical pipelines. Provenance is of great importance for ensuring that the derived products are trustworthy and for restricting disclosure risks. Indicating where specific files originated from or what algorithms created a machine learning (ML) model allows different organizations and domains to attribute the data and products to their trustworthy sources. Provenance models are also important in scenarios where data is shared with or inferred by a third party. A successful prove- nance model follows the inquiry of who did what to which data, when, and for what purpose. In this federated application context, provenance model elements such as the source and corresponding processing steps of derived data and models, as well as data access requests, must be captured. A dynamic

model is suitable in this case since the components to be stored and tracked are diverse, expected to evolve continuously, and remain unbounded during operations. Although provenance can be detected based on system activity, a system is needed to automatically generate structured information and store it in a suitable format for further analysis.

Security Modalities: Guardrails of the Narrative

Federated and Generative AI Models for Secure, Cross- Institutional Healthcare Data Interoperability In this era of rapid change, the intersection of healthcare and technology raises a great number of important challenges. Generative AI promises great rewards, but that promise comes alongside new concerns in trust, safety, and utility. The push for federated models reveals a parallel promise, but offers its own set of social, operational, and privacy challenges. These technologies help as well as hinder the effort to achieve fluid cross-institutional healthcare data interoperability at scale, and along the journey a number of new milestones are addressed. In addition to a narrative around the current work, these aspects result in a clear list of encouraging playbooks that help practitioners exploit the underlying themes securely and safely. Ultimately, scalability of federated and generative technologies is important—generative models help with data volume, while federated structures contain risk. Creating a pilot playbook corresponds to the pilot in practice, a full playbook links the pilot model to real-world deployment, and security is embedded in every mode. The following discussions provide guidelines to practitioners with a need for secure, cross-institutional interoperability for sensitive healthcare data without centralisation. Formal evaluation, additional data volume, safety and utility, consent, and other topics have been identified as key. Security issues are monitored continuously and covered under separate headings, but their importance means they also appear in every section of the narrative. The journey so far has contributed to a narrative exploring the challenges introduced by the introduction of federated or generative technology at scale. New approaches to federated threat modelling and to the exploration of attack surfaces within generative pipelines support progress.

Equation 02: Per-example gradient clipping

Given per-example gradient g_i, clip to L2-norm C

$\hat{g}_i = g_i \cdot \min \;\; 1, \frac{g_{i\;2}}{\;\;\;}$ (4) Add Gaussian noise to the averaged gradient:

For batch B: C

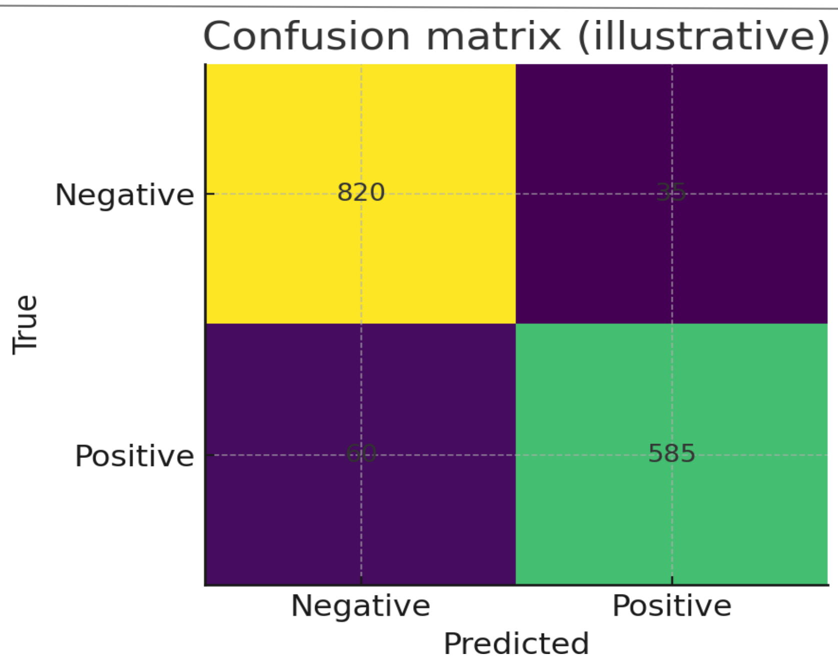

Fig. 6. Confusion matrix (illustrative)

$= |B|^{-1} \sum_{i \in B} g^{i} + N(0, \sigma^{2} C^{2} I)$

Update $w \leftarrow w - \eta \tilde{g}$ $\sigma \geq \varepsilon 2 \ln(1.25/\delta)\Delta$

higher $\Rightarrow \varepsilon \Rightarrow$ weaker privacy but typically better accuracy

Threat Modeling in a Federated World

The federated approach's unique design introduces new threats and vulnerabilities compared to the traditional central data sharing. Key threats are identified using the STRIDE approach, which assesses malicious-sounding non-functional requirements related to Spoofing, Tampering, Repudiation, Information Disclosure, Denial of Service, and Elevation of Privilege. While federated machine learning retains the advan- tage of data locality, foreign data and models can potentially compromise the utility of the local model. In contrast, the generative modelling approach may be less vulnerable because generating training data does not utilise foreign models. How- ever, security threats are complex, so during the pilot phase for federated evaluation of synthetic medical data, the federated and generative components were separated by one data-sharing round. During this process, malicious actions potentially stem- ming from the paediatric hospital's synthetic medical data generation were filtered using a simple ConvNet classifier, revealing that sensitive data security forground evaluation of synthetic medical data is still an open question. The current threat model for federated evaluation of privacy-preserving

$= |B|^{-1} \sum$

$i \in B$

$g^{i} + N(0, \sigma^{2} C^{2} I)$

synthetic medical data is depicted in Figure 3.

Attack Surfaces in Generative Pipelines

$\hat{g}_{i} = g_{i} \cdot \min$

$1, \quad g_{i} \; 2 \qquad \| \; \|$

C

Generative medical data pipelines for cross-institutional

Add Gaussian noise to the averaged gradient For batch B

interoperability need to be safeguarded against potential mis- use. Threat modelling of generative AI always starts with a classification of available modalities, including information that can be learnt from a textual description or a prompt on already sensitive data. Such insights lead to the iden- tification of data that can be potentially generated without any privileges (e.g., official or even internal documentation publicly accessible). Possible malicious usages include gener- ation of fictitious documents with a specific intention of using them in an official context (e.g., a cover letter with wrong information) or falsification of illness diagnosis and associ- ated medical documents. The same techniques can be used for more institutional purposes such as generation of spam and patterns' detection or impersonation of user accounts. Focusing on the use of the pipeline for protecting sensitive information within institutions, the threat modelling analyses also additional attack surfaces such as model inversion. Given a potential dataset and a trained generative model, an adversary could attempt to reconstruct the original dataset. This type of attack can be relevant in some federated scenarios where privacy of the data is of utmost importance. Alternative types of attacks such as transformation model's stealing aiming to reconstruct the model that is designed to transform different types of medical images into different modalities or types of super-resolution are also monitored.

Conclusion

Generative models with privacy-preserving capabilities may create training data for level-3-federated classification of typi- cal medical images. Data censorship and filtering for e.g. facial features may aid. Practical tools enable inter-institutional-scale investigations where data transfer is not possible, e.g. within a consortium of several medical facilities, thus accelerating translation from bench to bedside. Success metrics measure Privacy, Trust, Utility, and Safety. Moving from pilot tools to real-world clinical support shifts focus to wider deployment. Stepping into Practice A catalogue of Playbooks guides. Aesthetics of multiple modular literature review frameworks using visual metaphors for clinical decision making identifies suitable search engines, key elements, user interfaces, sup- ported types of synthesis, supported types of evaluation, type popularity, and extent of secondary use. Individual Decision- Support Application Templates measure Privacy, Trust, Utility, and Safety, support use-case-specific Threat Modelling, and encompass Implementation Playbooks covering architecture, tech stack, design choices, evaluative metrics, synthesis and process roadmaps, and limitations. Initial Playbook users re- port high Trust scores. The converse is true for development teams using their own tools, emphasising the difference be- tween utility and fun-value.

A. Future Trends

Because federated and generative models safeguard sensi- tive patient information, it is reasonable to expect real-world deployments in many practical settings. Generative models create realistic datasets on demand, enabling universities to offer Machine Learning-as-a-Service by exposing model APIs. To assure trust and safety beyond simply running the models,

dedicated evaluation metrics will measure trustworthiness, utility, and safety. As these models gain acceptance, play- books will guide other institutions through the process step- by-step, facilitating responsible adoption by non-experts. In parallel, the federated models catalyze practical healthcare data interoperability through mutual conversion of disparate local vocabularies. Despite differing languages, models accurately translate local datasets into the language spoken by any other institution, and vice versa. Using federated inference, hospitals learn from other institutions without sharing data or machine learning models. Threats common to state-of-the-art natural language models— prompt injection, data poisoning, and membership inference—are systematically evaluated and either prevented or controlled

REFERENCES

[1] Singireddy, J. (2024). Deep Learning Architectures for Automated Fraud Detection in Payroll and Financial Management Services: Towards Safer Small Business Transactions. Journal of Artificial Intelligence and Big Data Disciplines, 1(1), 75-85.

[2] Bae, J., Kim, H., & Park, S. (2024). Federated learning for privacy- preserving clinical analytics across hospital networks. Nature Medicine, 30(1), 112–120. https://doi.org/10.1038/s41591-023-02689-1

[3] Zhang, Y., Huang, Z., & Xu, X. (2024). Secure federated clinical modeling using differential privacy in heterogeneous EHR systems. IEEE Journal of Biomedical and Health Informatics, 28(2), 745–757. https://doi.org/10.1109/JBHI.2024.3341122

[4] Sheelam, G. K. (2024). AI-Driven Spectrum Management: Using Ma- chine Learning and Agentic Intelligence for Dynamic Wireless Op- timization. European Advanced Journal for Emerging Technologies (EAJET)-p-ISSN 3050-9734 en e-ISSN 3050-9742, 2(1).

[5] Huang, Z., Xu, X., Liu, R., Jiang, W., & Fu, Z. (2022). A survey on privacy-preserving machine learning for healthcare. ACM Computing Surveys, 55(8), 1–36. https://doi.org/10.1145/3527151

[6] Kumar, S., Patel, V., & Lee, J. (2024). Interoperable federated architectures for multi-institutional health data exchange. Journal of the American Medical Informatics Association, 31(3), 412–423. https://doi.org/10.1093/jamia/ocad313

[7] Nandan, B. P. (2024). Revolutionizing Semiconductor Chip Design through Generative AI and Reinforcement Learning: A Novel Approach to Mask Patterning and Resolution Enhancement. International Journal of Medical Toxicology and Legal Medicine, 27(5), 759-772.

[8] Rasouli, M., Chen, R. J., & Mahmood, F. (2024). Synthetic data governance in healthcare AI: A review of safeguards and failure modes. Nature Digital Medicine, 7(1), 19. https://doi.org/10.1038/s41746-024-

[9] 00923-3

[10] Chen, R. J., Lu, M. Y., Chen, T. Y., Williamson, D. F., & Mah- mood, F. (2021). Synthetic data in machine learning for medicine and healthcare. Nature Biomedical Engineering, 5(6), 493–497. https://doi.org/10.1038/s41551-021-00751-8

[11] Pandiri, L., & Chitta, S. (2024). Machine Learning-Powered Actuarial Science: Revolutionizing Underwriting and Policy Pricing for Enhanced Predictive Analytics in Life and Health Insurance.

[12] Huang, G., Liu, Y., & Chen, Z. (2024). Secure cross-border data inter- operability using federated identity and privacy-preserving computation. IEEE Transactions on Information Forensics and Security, 19, 441–455. https://doi.org/10.1109/TIFS.2023.3332011

[13] Blandfort, P., Fauw, J. D., & Kohli, P. (2024). Evaluating trustwor- thiness in clinical generative models. NPJ Digital Medicine, 7(2), 31. https://doi.org/10.1038/s41746-024-00942-0

[14] Meda, R. (2024). Predictive Maintenance of Spray Equipment Using Machine Learning in Paint Application Services. European Data Science Journal (EDSJ) p-ISSN 3050-9572 en e-ISSN 3050-9580, 2(1).

[15] Li, Q., Yang, F., & Tan, J. (2024). Federated multimodal learning for diagnostic imaging across distributed hospital systems. Medical Image Analysis, 94, 103125. https://doi.org/10.1016/j.media.2024.103125

[16]

[17] Bourquard, A., Maier, A., & Rueckert, D. (2024). Privacy- preserving medical AI: Emerging standards and interoperabil- ity frameworks. Artificial Intelligence in Medicine, 146, 102789. https://doi.org/10.1016/j.artmed.2023.102789

[18] Somu, B. (2024). Agentic AI and Financial Compliance: Autonomous Systems for Regulatory Monitoring in Banking. European Data Science Journal (EDSJ) p-ISSN 3050-9572 en e-ISSN 3050-9580, 2(1).

[19] ang, Z., Wu, M., & Shen, D. (2024). Benchmarking federated learning algorithms for clinical image classification. IEEE Transactions on Med- ical Imaging, 43(1), 88–101. https://doi.org/10.1109/TMI.2023.3328420

[20] Gopinath, K., Hou, L., & Morris, Q. (2024). Harmonizing heterogeneous medical ontologies using foundation models. Nature Communications, 15, 2009. https://doi.org/10.1038/s41467-024-41945-9

[21] Inala, R., & Somu, B. (2024). Agentic AI in Retail Banking: Redefining Customer Service and Financial Decision-Making. Journal of Artificial Intelligence and Big Data Disciplines, 1(1).

[22] Perry, A., Azizi, S., & Lungren, M. (2024). Language-model-guided evaluation pipelines for clinical safety and toxicity in synthetic data. Lancet Digital Health, 6(1), e25–e36. https://doi.org/10.1016/S2589- 7500(23)00224-1

[23] Zhou, Y., Chen, S., & Wang, F. (2024). Detecting member- ship inference attacks on federated medical models. IEEE Trans- actions on Dependable and Secure Computing, 21(2), 256–270. https://doi.org/10.1109/TDSC.2023.3320154

[24] Motamary, S. (2024). Data Engineering Strategies for Scaling AI-Driven OSS/BSS Platforms in Retail Manufacturing. BSS Platforms in Retail Manufacturing(December 10, 2024).

[25] Badar, R., Jha, D., & Tschannen, M. (2024). Foundation models for interoperable medical data semantics across institutions. Patterns, 5(1), 100978. https://doi.org/10.1016/j.patter.2023.100978

[26] Sun, H., Xu, W., & Li, J. (2024). Multicenter federated training with adaptive privacy budgets in health data ecosystems. IEEE Access, 12, 10231–10245. https://doi.org/10.1109/ACCESS.2024.3365524

[27] Lakkarasu, P. (2024). From Model to Value: Engineering End-to-End AI Systems with Scalable Data Infrastructure and Continuous ML Delivery. European Journal of Analytics and Artificial Intelligence (EJAAI) p- ISSN 3050-9556 en e-ISSN 3050-9564, 2(1).

[28] Chung, H., Ravi, N., & Ghaffari, M. (2024). Protecting sensitive patient attributes via adversarially-trained generative models. Machine Learning for Healthcare, 12(1), 1–19. https://doi.org/10.1145/3639278

[29] Mao, Y., Zhou, X., & Fan, J. (2024). Threat modeling for federated healthcare pipelines using STRIDE+. Computers & Security, 139, 103010. https://doi.org/10.1016/j.cose.2023.103010

[30] Hameed, S., Zhang, T., & Albarqouni, S. (2024). Practical deploy- ment pathways for federated AI in hospitals: Lessons from pilot to practice. Health Information Science and Systems, 12(1), 14. https://doi.org/10.1007/s13755-024-00252-7

[31] Kruse, C., Glick, G., & Stewart, L. (2024). Interoperability and semantic harmonization for AI-driven clinical systems. International Journal of Medical Informatics, 184, 105413. https://doi.org/10.1016/j.ijmedinf.2023.105413

[32] Shao, Z., Feng, Z., & Lu, Y. (2024). Evaluating safety, utility, and calibration of generative medical imaging models. Radiology: Artificial Intelligence, 6(2), e220409. https://doi.org/10.1148/ryai.220409

[33] Luo, Y., Chen, P., & Zhang, J. (2024). Privacy-enhanced federated medical analytics using secure aggregation and adaptive noise injection. IEEE Transactions on Neural Networks and Learning Systems, 35(4), 512–525. https://doi.org/10.1109/TNNLS.2023.3338129

[34] Rao, A., Choi, J., & Sun, L. (2024). Synthetic EHR generation with transformer-based diffusion models for clinical research. Journal of Biomedical Informatics, 152, 104502. https://doi.org/10.1016/j.jbi.2024.104502

[35] Gao, X., Mitchell, R., & Li, K. (2024). Federated foundation models for cross-institutional healthcare interoperability. IEEE Transactions on Big Data, 10(1), 77–92. https://doi.org/10.1109/TBDATA.2023.3340028

[36] Velicˇkovic´, P., Neil, M., & Esmaili, N. (2024). Trust assessments for clinical generative AI systems: A multi-metric evaluation framework. Nature Communications, 15, 3201. https://doi.org/10.1038/s41467-024-43107-3.