Secure distributed machine learning in healthcare: a study on FAIR, compliance and cybersecurity for federated learning

Plug, R.B.F.

1 Introduction

> *The future of artificial intelligence should be shaped by the values of human compassion, ethics, and scientific rigor.*
>
> **— Fei-Fei Li**
> *On the nature of ethics in artificial intelligence.*

1.1 Research Context

Recent breakthroughs in artificial intelligence provide new opportunities for medical research, especially in real-time surveillance and modelling of infectious disease dynamics, thereby improving patient outcomes and making healthcare delivery more efficient. However, the exponential growth of healthcare data opens up new challenges regarding to how this data is collected, processed, shared, and analysed [1]. This brings us to discussions around *data ownership*: who actually owns patient data and who can legally and ethically use it, and in which cases? We find challenges in *data interoperability*: how can we make sure data from different healthcare facilities and research programmes are compatible with each other? If we want to share and use data, we run into issues surrounding *regulatory compliance*, which is a very complex legal topic for protected data categories such as healthcare data. Finally we also consider *cybersecurity*, how do we make sure that our data stay secure and third parties cannot get unauthorised access to our sensitive data and studies?

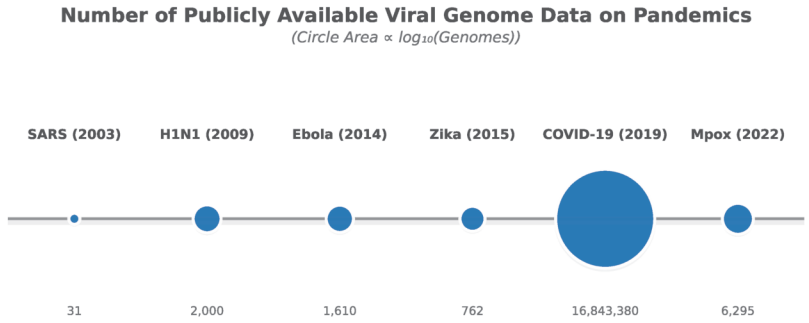

Fig. 1.1.: Publicly shared full-genome sequences for six major zoonotic epidemics and pandemics. The circle area scales with $\log_{10}$ of the number of genomes. Counts are shown below each point. [5, 6, 7, 8, 9, 10]

All these concerns together shape research and implementation gaps that prevents us from providing the most effective, equitable, sustainable and safe use of health data in artificial intelligence applications. The implications of these gaps are particularly pronounced in resource-constrained regions where populations are often excluded from leading studies and healthcare applications. This is compounded by the lack of structured processes around health data management and data stewardship, which ultimately impede the chain of diagnostics and medical decision making, especially between local clinics and regional hospitals, a gap that becomes critical when rapid epidemiological action is required during outbreaks such as the recent pandemic. This ultimately leads to suboptimal health outcomes for a significant part of the world population [2, 3, 4].

Healthcare Data in Numbers

Healthcare data generation has increased dramatically, driven by increasing digitalisation of patient files in electronic health records (EHRs), data-heavy diagnostics such as medical imaging and increasing use of wearable devices to track patient biometrics.

Healthcare data is approximated to account for about 30% of the world's total data volume [11]. In 2020, the global healthcare data volume reached over 2 000 exabytes, which is over 250 gigabytes of data per person on earth. A figure which is expected to multiply almost six fold, surpassing a total of 11 000 exabytes by 2030 [12]. Not only has healthcare data exponentially grown in size, but healthcare data is also tracked in more detail and over longer continuous periods. Longitudinal surveillance and genomic datasets for infectious diseases such as H1N1 and SARS-CoV-2 now constitute some of the largest continuously growing sources of structured epidemiological data as shown in Figure 1.1.

To harness the potential of the volume and veracity of the data being produced, the European Commission has launched the European Health Data Space (EHDS) initiative in 2022 as part of the movement towards universal data spaces. The EHDS is aimed to increase the findability, accessibility, interoperability and reusability (FAIR) of health data across EU member states, while maintaining compliance with data privacy and protection regulations such as the GDPR. [13].

This is a large contrast to the situation in resource-constrained environments such as Africa. Data infrastructures in Africa remain fragmented and often rely on external service providers for their health information systems, which do not guarantee data locality. With an exponentially growing population, it is unclear exactly how much data is being processed, how much data is exposed to privacy and security risks and how much data is being lost due to a lack of digitisation. Despite increasing digital connectivity, with Sub-Saharan Africa reaching 40% of the population having internet connectivity [14], the majority of the African healthcare institutions still operate primarily on paper-based systems [15].

The studies performed in this dissertation are performed in the context of the Virus Outbreak Data Network (VODAN-Africa). VODAN's focus on real-world clinical viral outbreak and outpatient data registration positions the research in the epidemiological data sciences. VODAN addresses the many challenges of health data in Africa in particular, which poses additional challenges due to resource shortages and lack of unified healthcare and data policies [4]. Within VODAN research was performed on the digitisation of epidemiological data to a FAIR-format by design and the localisation of clinical data generation and management in-residence. This is following a federated data structure, which can better support the sheer scale that African healthcare needs will grow towards in the coming decades. [2].

The disparity in healthcare data management is partly driven by data inequity. Traditionally, research involving African data has involved extracting and processing data outside the continent, often without direct benefit for local communities and scientific communities. This is a practice known as parachute research [16]. Addressing these disparities requires strategies emphasizing data locality and ownership, ensuring data remains at the source yet can be used for scientific study and real-time epidemic surveillance. This thesis explores approaches that bridge this research gap, enabling equitable data management and value creation at local levels through federated data methods, which also address the issue of scale as mentioned.

Healthcare Data Reuse

Lack of data interoperability between healthcare institutions and between research projects is an ever ongoing problem that reduces the chances of reusing past data. This is especially problematic in cases where there are data of rare diseases or one-off events such as a global pandemic. This is not only problematic to research potential, but also increases the burden on research funding. In

the EU alone, lacking data interoperability and reusability is estimated to cost healthcare providers approximately €10 billion annually [17]. If we consider the total costs of redundant data collection, duplicated data processing, missed research opportunities and delayed treatments cost estimates can go as high as €26 billion per year [17]. Looking towards the health data ecosystems across Africa in VODAN, we see that healthcare information systems are subject to the concept of digital black hole [18].

This is characterised by geographically isolated, unconnected and non-interoperable health data typically stuck in paper format or legacy systems. This leads to the problem of missing data as Ebola or COVID-19 [4], where we simply do not have the necessary volume and quality of data to compare or analyse the local situation due to problems in findability and accessibility, ultimately leading to non-reusability. Approximately 40% of clinical data captured globally is never reused for secondary purposes due to such interoperability barriers and lacking findability and interoperability due to inadequate metadata [17].

To resolve challenges around interoperability and ultimately reusability of data, we require standardisation of data management practices. An existing example of global data harmonisation is the WHO Global Influenza Surveillance and Response System (GISRS) [19], which has collected and standardised virological metadata in a consistent schema since the 1950s. This type of standardisation requires manual data engineering and does not scale well beyond one use case given the variety of medical specialisations, data types and standards across global clinics. One of the most important proposals surrounding this topic, and one that is central to this dissertation, are the FAIR principles. These are a set of principles that define standards around making data findable, accessible, interoperable, and reusable [20] through rich metadata annotations. One of the advanced topics of interest

there is FAIR data creation directly at the point-of-care. This requires the utilisation of data forms based on semantic metadata templates. This can guarantee a certain level of interoperability and reusability for all data created using that methodology. The Personal Health Train (PHT) paradigm operationalises these FAIR data by enabling federated data analyses on FAIR data. Here analyical algorithms such as federated learning models visit data and perform secure computations at their origin without moving sensitive data across different locations [21]. The use of FAIR and federated data methodologies are one of the main issues addressed in the articles presented in this dissertation.

AI Ethics in Healthcare

The increasing use of AI techniques such as machine learning has sparked discussions around ethics and safety, especially in domains using sensitive private information like healthcare. According to a recent study, European healthcare providers expressed concerns about the ethics and GDPR compliance of using AI-driven healthcare solutions [22], especially those that profile patients for communicable-disease risk stratification. GDPR explicitly emphasises patient rights, including consent, data minimisation, and purpose limitation. Together with the recent AI-act, this makes it a legal and ethical necessity to consider compliance and safety when deploying AI solutions in healthcare [23]. In global cross-country settings like those found in VODAN-Africa, ethical compliance is further complicated by the diverse range of local regulations and ethics. This makes it necessary to find approaches that can adapt across multiple geographies and jurisdictions.

The central technique we study to provide compliance, privacy-preservation and security in machine learning is federated learning. This technique addresses ethical and regulatory considerations by allowing shared

training of a model on localised datasets without sharing actual health data. This approach naturally aligns with GDPR's emphasis on data minimisation and local compliance, as personal health data never leaves the point of care [24], while still providing opportunities for data reuse.

Additionally, federated learning enables local value creation and provides strong guarantees for data ownership. This ultimately means that everyone that contributes, can exactly determine what data they allow to be used without exposure, while equally benefitting from the resulting collaboratively trained model. In this dissertation federated learning is investigated, with a research focus on state-of-the-art privacy-preservation and security. We model and verify how secure federated learning methods can be applied to privacy-sensitive healthcare analytics in diverse regulatory landscapes by utilising security and computation sharing techniques from adjacent machine learning methods such as cryptography and secret sharing from secure multi-party computation.

Cybersecurity in Healthcare

With the massive volume of healthcare data and the sensitive nature of such data, healthcare data has become an attractive target for cybercriminals. With the massive volume of healthcare data and the sensitive nature of such data, healthcare data has become an attractive target for cybercriminals. Attacks and eventual breaches on healthcare institutions have been increasing dramatically in recent years. In 2022 alone, healthcare organisations globally experienced an average of 1 410 cybersecurity threats per week, a 74% increase compared to the previous year [25]. Within the European Union, the average healthcare data breach cost in 2023 has been estimated around €300 thousand [26], whereas the average cost of a successful healthcare data breach in the US reached $10.93 million in 2022, representing the

highest breach cost of any industry [27]. Such breaches not only compromise patient confidentiality, but also threaten to disrupt day-to-day clinical operations. The increased use of AI in clinical settings provides another threat vector for malign actors to exploit if given the chance.

The method of federated learning that we study provides privacy-preserving properties by design, but remains vulnerable to sophisticated cybersecurity threats. These are threats that may involve compromised clients and reconstruction of source data which may expose private data that the method seeks to preserve. Mitigating these risks is non-trivial and requires newly developed methods to process data, new machine learning training methodologies and measures to ensure no single party can compromise security. Adding additional layers of security to federated learning may have an impact on model performance and efficiency, and the exact nature of that impact has to be validated through empirical measurements. In this dissertation we perform an ablation study to evaluate our secure distribution machine learning framework based on federated learning. With this foundational evidence is provided for practical application of secure federated learning for healthcare use cases.

The central thesis of this dissertation is that the proposed Secure Distributed Machine Learning (SDML) methodology is a scalable and GDPR-compliant framework for privacy-preserving machine learning across distributed healthcare use-cases. SDML incorporates the principles of regulatory compliance and cryptographic security into one design-science-based architecture. The subsequent chapters collectively build and empirically substantiate this framework, showing that secure federated models can be built specifically for healthcare settings without requiring centralisation of data, even in challenging compute-restricted environments.

This dissertation addresses each of the key issues we have mentioned by systematically exploring the potential and use of federated learning-based artificial intelligence in a secure and privacy-preserving manner. The focus is on the topics of FAIR, compliance and cyber security. By examining these domains through both theoretical desk study and empirical benchmark analyses, this research contributes to a nuanced understanding of how distributed AI applications in healthcare can be designed and utilised in practice. This research aligns that goal with considerations about global ethical standards, regulatory frameworks, and the practicalities of cross-border clinical use cases [28, 29, 30]. In this dissertation I seek to cover existing research gaps in both FAIR data and secure machine learning, with the aim to further secure, responsible and equitable use of health data. Ultimately to increase opportunities for clinical research and improve patient outcomes globally, particularly within underserved communities.

1.2 Research Outline

The main body of research in this dissertation is divided into four distinct chapters, each building upon the previous chapter. This work is structured using a cumulative research approach following the Design Science Research (DSR) framework.

The dissertation starts out with a broad literature review in **chapter 2** on a FAIR data framework in the resource-constrained, cross-country African healthcare context. This chapter also serves as a key piece of research in which specific research gaps surrounding FAIR data and federated data infrastructures are identified. This chapter provides the knowledge and artifacts that we need to design a viable solution.

In the following two chapters we perform the design steps of our research framework. In **chapter 3** we investigate and design the potential of a data management framework in healthcare data analytics

that provides GDPR compliance and FAIR conformity by design. This chapter addresses gaps in data interoperability and regulatory compliance. This is followed by **chapter 4**, which builds upon the previous chapters to do a technical deep-dive into cybersecurity methodologies for designing a secure version of federated learning that has potential for resource-constrained healthcare environments. These two chapters together conclude the design cycle of our DSR approach.

Finally, in **chapter 5** we seek to empirically validate the approach we have built towards in the previous chapters. In this chapter we perform a simulation study to benchmark an implementation of our secure federated learning solution. We replicate a real use case by using industry-standard open clinical benchmark data and by constraining available resources. We analyse the results to check for non-inferiority in model performance and use ablation to measure the impact on model training runtime.

References

[1] Andre Esteva, Alexandre Robicquet, Bharath Ramsundar, et al. „A guide to deep learning in healthcare“. In: *Nature Medicine* 25.1 (2019), pp. 24–29 (cit. on p. 5).

[2] Alicia Martínez-García, Celia Alvarez-Romero, Esther Román-Villarán, Máximo Bernabeu-Wittel, and Carlos Luis Parra-Calderón. „FAIR principles to improve the impact on health research management outcomes“. In: *Heliyon* 9.5 (2023), e15733. ISSN: 2405-8440 (cit. on pp. 6, 8).

[3] Núria Queralt-Rosinach, Rajaram Kaliyaperumal, César H. Bernabé, et al. „Applying the FAIR principles to data in a hospital: challenges and opportunities in a pandemic“. In: *Journal of Biomedical Semantics* 13.12 (2022) (cit. on p. 6).

[4] Mirjam van Reisen, Mia Stokmans, Munyaradzi Mawere, et al. „FAIR practices in Africa“. In: *Data Intelligence* 2.1–2 (2020), pp. 246–256 (cit. on pp. 6, 8, 9).

[5] Jianfei Hu, Jing Wang, Jing Xu, et al. „Evolution and Variation of the SARS-CoV Genome“. In: *Genomics, Proteomics & Bioinformatics* 1.3 (2003), pp. 216–225 (cit. on p. 6).

[6] Yvonne C.F. Su, Justin Bahl, Udayan Joseph, et al. „Phylodynamics of H1N1/2009 influenza reveals the transition from host adaptation to immune - driven selection“. In: *Nature Communications* 6 (2015), p. 7952 (cit. on p. 6).

[7] Gytis Dudas, Luiz Max Carvalho, Trevor Bedford, et al. „Virus genomes reveal factors that spread and sustained the Ebola epidemic“. In: *Nature* 544.7650 (2017), pp. 309–315 (cit. on p. 6).

[8] Sofia G. Seabra, Pieter J. K. Libin, Kristof Theys, et al. „Genome-wide diversity of Zika virus: Exploring spatio-temporal dynamics to guide a new nomenclature proposal“. In: *Virus Evolution* 8.1 (2022), veac029 (cit. on p. 6).

[9] GISAID Initiative. *hCoV-19 EpiCoV Global Submission Tracker*. 219 countries and territories have shared 16,843,380 SARS-CoV-2 genomes since 10 January 2020. 2025 (cit. on p. 6).

[10] GISAID Initiative. *hMpxV EpiPox Global Submission Tracker*. 52 countries have shared 6,295 mpox virus genomes since 1 January 2022. 2025 (cit. on p. 6).

[11] Jane Thomason. „Data, digital worlds, and the avatarization of health care“. In: *Global Health Journal* 8.1 (2024), pp. 1–3. ISSN: 2414-6447 (cit. on p. 7).

[12] Klaus Boehncke, Guillaume Duparc, Jonathan Sparey, and Andre Valente. *Tapping Into New Potential: Realising the Value of Data in the Healthcare Sector*. Accessed: April 7, 2025. 2023 (cit. on p. 7).

[13] Rada Hussein, Lucas Scherdel, Frederic Nicolet, and Fernando Martin-Sanchez. „Towards the European Health Data Space (EHDS) ecosystem: A survey research on future health data scenarios“. In: *International Journal of Medical Informatics* 170 (2023), p. 104949. ISSN: 1386-5056 (cit. on p. 7).

[14] Hanim Maria Astuti and Lateef Adeshina Ayinde. „Uneven Progress: Analyzing the Factors Behind Digital Technology Adoption Rates in Sub-Saharan Africa (SSA)“. In: *Data Policy* 7 (2025), e23 (cit. on p. 7).

[15] Beatrice Kuvuna, Moriasi Nyanchoka, Fatuma Guleid, et al. „Community-Based Health Information Systems in Africa: A Scoping Review of Data Generation, Utilization, and Community Empowerment“. In: *Wellcome Open Research* 9 (2024), p. 485 (cit. on p. 7).

[16] I. P. J. Pocock, Andrew T. Knight, Nicola J. van Wilgen, et al. „From parachuting to partnership: Fostering collaborative research in protected areas“. In: *Journal of Applied Ecology* (2024) (cit. on p. 8).

[17] European Commission: Directorate-General for Research and Innovation and PwC EU Services. *Cost-benefit analysis for FAIR research data – Cost of not having FAIR research data*. 2018 (cit. on p. 9).

[18] Samson Yohannes Amare, Araya Abrha Medhanyie, and Mirjam van Reisen. „Data Visiting in Digital Black Holes: FAIR Based Digital Health Innovation during War“. In: *Tigray. War in a Digital Black Hole. Book 3*. Final draft. Langaa RPCIG, 2024, pp. 477–508. ISBN: 9789956554188 (cit. on p. 9).

[19] Alan J. Hay and John W. McCauley. „The WHO global influenza surveillance and response system (GISRS)—A future perspective“. In: *Influenza and Other Respiratory Viruses* 12.5 (2018). Epub 2018 Jun 25, pp. 551–557 (cit. on p. 9).

[20] Mark D. Wilkinson, Michel Dumontier, IJsbrand Jan Aalbersberg, et al. „The FAIR Guiding Principles for scientific data management and stewardship“. In: *Scientific Data* 3 (2016), p. 160018 (cit. on p. 9).

[21] Oya Beyan, Ananya Choudhury, Johan van Soest, et al. „Distributed Analytics on Sensitive Medical Data: The Personal Health Train“. In: *Data Intelligence* 2.1-2 (2020), pp. 96–107 (cit. on p. 10).

[22] Rocco de Filippis, Abdullah Al Foysal, Vincenzo Rocco, et al. „The risk perspective of AI in healthcare: GDPR and GELSI framework (Governance, Ethical, Legal and Social Implications) and the new European AI Act“. In: *Italian Journal of Psychiatry* 10.1 (2024) (cit. on p. 10).

[23] European Parliament and Council of the European Union. *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)*. `https://eur-lex.europa.eu/eli/reg/2016/679/oj`. Official Journal of the European Union, L 119, 4 May 2016, pp. 1–88. 2016 (cit. on p. 10).

[24] Nicola Rieke, Jonny Hancox, Wenqi Li, et al. „The Future of Digital Health with Federated Learning“. In: *npj Digital Medicine* 3 (2020), p. 119 (cit. on p. 11).

[25] Check Point Research. *Healthcare Cyberattacks Rising*. Accessed: 2025-04-07. 2022 (cit. on p. 11).

[26] European Union Agency for Cybersecurity (ENISA). *ENISA Threat Landscape: Health Sector*. Tech. rep. European Union Agency for Cybersecurity, 2023 (cit. on p. 11).

[27] IBM Security. *Cost of a Data Breach Report 2023*. Tech. rep. IBM Corporation, 2023 (cit. on p. 12).

[28] Rosie Richards. „Barriers on cross-border sharing of health data for secondary use and options to overcome these“. In: *The European Journal of Public Health* 32.Suppl 3 (2022), ckac129.367 (cit. on p. 13).

[29] Hui Yun Chan, Hui Jin Toh, and Tamra Lysaght. „Cross-jurisdictional Data Transfer in Health Research: Stakeholder Perceptions on the Role of Law“. In: *Asian Bioethics Review* 16 (2024), pp. 663–682 (cit. on p. 13).

[30] Miroslav Puskaric, Balasubramanian Chandramouli, Thomas Osmo, et al. „Privacy-Preserving Workflow for the Cross-Border Federated Analysis of Clinical Data“. In: *Studies in Health Technology and Informatics* 316 (2024), pp. 1637–1641 (cit. on p. 13).