CHAPTER

Recent advances in federated learning for digital healthcare systems

4

Pooja Mohnani[1], Christoph Thümmler[2], Angelica Avila Castillo[2], Rasha Tolba[2], Alessandro Bassi[1], Antoine Simon[3], Anastasius Gavras[1], Orazio Toscano[4] and Pascal Haigron[3]

[1]*Eurescom GmbH, Heidelberg, Germany*
[2]*6G Health Institute GmbH, Markkleeberg, Germany*
[3]*Univ Rennes, Inserm, LTSI - UMR 1099, Rennes, France*
[4]*Ericsson, Genova, Italy*

4.1 Introduction

In traditional healthcare systems, sharing and analyzing patient data for research, diagnosis, and treatment is hampered by privacy concerns and regulatory restrictions. Federated healthcare platforms enable the deployment of federated learning (FL), a novel framework to overcome these barriers by allowing healthcare organizations to collaborate without sharing sensitive patient information (Li, Wang, et al., 2020; McMahan, Ramage, et al., 2017; Yang et al., 2019). Instead, each organization retains control of its data, and only aggregated statistics or model updates are shared between organizations. This distributed approach ensures that patient privacy is protected, while still allowing for collaborative sharing, analysis, and model training.

The concept of federated healthcare platforms addresses the growing recognition of the value of healthcare data and the need for data-driven insights to improve patient care and outcomes. As shown in Fig. 4.1, by bringing together diverse datasets from multiple organizations, federated healthcare platforms enable the development of more accurate diagnostic models, personalized treatment recommendations, and real-time predictive analytics (Rieke et al., 2020). These platforms facilitate knowledge sharing and collaboration, ultimately improving the effectiveness and efficiency of healthcare delivery.

Thus this opens many research challenges and possibilities that our society can and must address to prepare for current and future innovation in healthcare. Through the implementation of FL and the adoption of secure, distributed healthcare data management practices, can unlock the full potential of digital healthcare systems, leading to improved patient outcomes and a more efficient healthcare ecosystem.

Federated Learning for Digital Healthcare Systems. DOI: https://doi.org/10.1016/B978-0-443-13897-3.00005-9

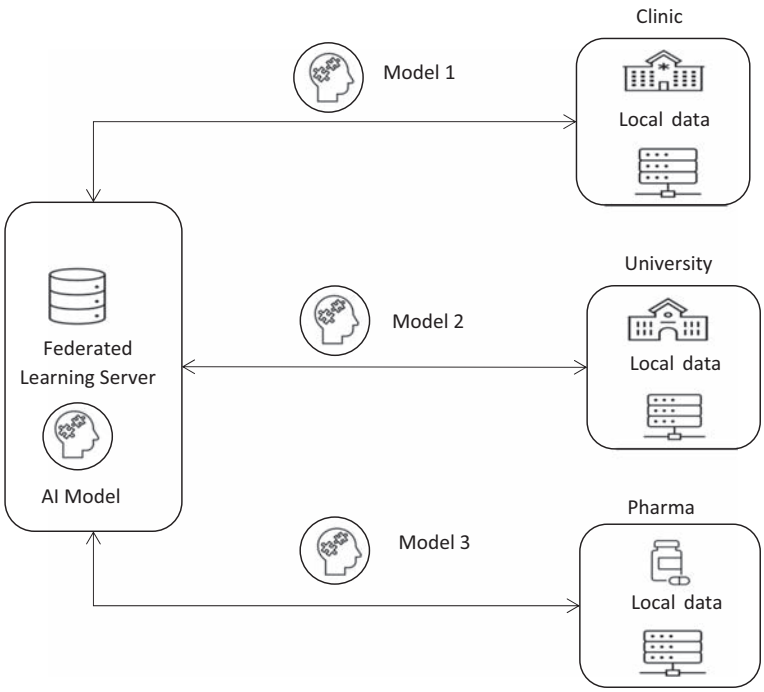

FIGURE 4.1

The concept of federated healthcare platforms.

4.1.1 Key contributions of the chapter

The following are the significant contributions of this chapter:

1. Identification of key FL application areas.
2. Highlighting crucial considerations essential for the development of FL systems, with a specific focus on ensuring scalable and reliable privacy preservation.
3. Addressing significant challenges and exploring the prospects of FL in healthcare, examining the anticipated evolution of these systems in the coming times.

4.1.2 Chapter organization

Section 4.1.2 presents the related works on federated learning and how it is evolving. Section 4.2 presents the perceptions of federated digital healthcare platforms in medical decision systems. Section 4.3 elaborates on the need and importance of privacy preservation, security, and ethics in healthcare data. Section 4.4 discusses the challenges, future research directions, trending

technologies, and recent advances. Section 4.5 presents the view on FL and its integration with future 6G mobile communications networks. Finally, Section 4.6 concludes the chapter.

4.2 Related works

FL enables healthcare records that are located across different institutions to be connected without revealing personal information. Thus researchers, doctors, and data scientists can harness these extensive datasets from multiple hospitals without centralizing the data in one place. This approach effectively resolves crucial challenges related to the access rights of heterogeneous data (Dasaradharami Reddy & Gadekallu, 2023). This is advantageous as it reduces data security and privacy concerns by maintaining local data stores, in comparison to centralized machine learning (ML) techniques that require datasets to reside on one server (Song et al., 2022).

Rieke et al. (2020) discuss the current FL efforts for digital health and their impact on stakeholders, clinicians, patients, hospitals and practices, researchers and artificial intelligence (AI) developers, healthcare providers, and manufacturers. FL enables healthcare and related professionals to tackle the challenges of building unbiased models from datasets with optimized utilization of time, effort, and cost. By training algorithms within a hospital's secure firewall and sharing only the models, FL effectively addresses data governance concerns and ensures the maintenance of data security (Rieke et al., 2020). FL is equipped to capture a wide range of data variables, facilitating the analysis of patients based on their age, sex, and demographic characteristics. For instance, by accessing electronic medical records, patients with similar characteristics (cardiac arrest, mortality, ICU stay, etc.) could be known, and the need for their hospitalization could be predicted (Huang et al., 2019).

FL provides AI developers with access to larger and more diverse datasets that better represent current patients. As a result, AI-based healthcare solutions will have the capability to scale globally on an unprecedented level (Dasaradharami Reddy & Gadekallu, 2023). FL can have a significant impact on a wide range of stakeholders, including clinicians, patients, hospitals, medical researchers, and healthcare providers (Dasaradharami Reddy & Gadekallu, 2023).

4.3 Perceptions of federated digital platforms and their use in healthcare

Perceptions of federated systems are influenced by privacy concerns, interoperability, and governance. While benefits such as increased privacy and community participation attract users, challenges such as fragmentation, scalability, and

trustworthiness may favor centralized systems. Addressing these perceptions can support the development and adoption of federated systems, creating a more decentralized and user-centric landscape.

4.3.1 Use of federated learning in medical image processing

FL has gained significant attention in medical image processing due to its ability to leverage decentralized data while preserving patient privacy, as shown in Fig. 4.2. In the real world, numerous sources of medical data, including magnetic resonance imaging (MRI), X-ray, positron emission tomography (PET), and computerized tomography (CT), provide doctors with vast volumes of information in many different medical applications (Rehman et al., 2020; Shen et al., 2017). ML, especially deep learning, has revolutionized the automatic analysis of these medical images with different applications, from exam or object classification (e.g., normal versus abnormal mammogram, benign or malignant tumor) to image registration (alignment of multiple images) or image segmentation (e.g., brain tumor delineation). However, as the models are typically trained on data from a single center, they face challenges in generalization, leading to a decline in performance when applied to data from another center.

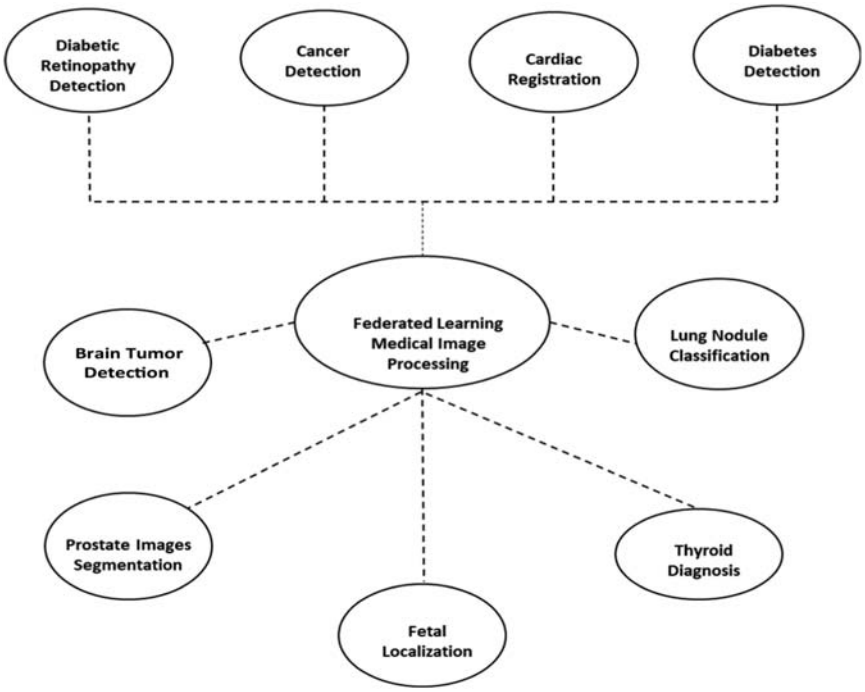

FIGURE 4.2

Applications of FL in medical image processing. *FL*, Federated learning.

FL holds immense potential in medical image analysis by facilitating collaborative training of ML models across healthcare institutions, ensuring adherence to privacy compliance. Li et al. (2020) demonstrated its feasibility in pneumonia detection using chest X-ray images from different hospitals. Zhang et al. (2021) demonstrated federated deep learning's ability to achieve COVID-19 detection accuracy while preserving patient data privacy. Different studies have demonstrated the effectiveness of FL in analyzing brain images. Abadi et al. (2016), and Silva et al. (2019) introduced this technique to aggregate encrypted updates without revealing sensitive information.

FL holds significant potential for medical image processing by enabling collaboration among healthcare institutions while ensuring data privacy. The referenced studies highlight the feasibility and effectiveness of FL in these domains, paving the way for future research and applications in healthcare.

4.3.2 Use of federated learning in Internet of Things-based smart healthcare applications

Recently, FL has emerged as a distributed collaborative AI approach, facilitating a range of intelligent Internet of Things (IoT) applications by enabling AI training on distributed IoT devices without the need for data exchange (Nguyen et al., 2021).

Fig. 4.3 illustrates a typical smart healthcare application based on FL. In this scenario, clinical data is collected from patients using onboard sensors. Multiple edge devices collaboratively execute the FL algorithm, and the resulting ML models assess the physical health of the patients. In urgent situations, the system

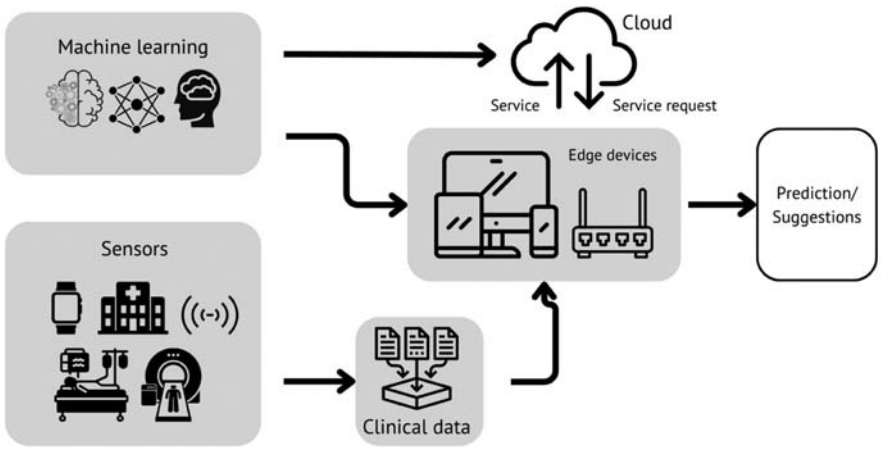

FIGURE 4.3

A typical FL-based smart healthcare application (Chang et al., 2021). *FL*, Federated learning.

can even request cloud-based emergency services. However, a limitation of traditional FL is its reliance on a reliable central server to aggregate model parameters uploaded by devices and distribute the global model to all participating devices (Chang et al., 2021).

In the domain of disease diagnosis and medical image processing, FL can leverage advancements in edge computing for enhanced benefits. Edge devices, such as smartphones and wearable devices, can participate in the FL process while preserving data privacy. Li et al. (2020) proposed a FL framework that leverages edge devices for the classification of electrocardiogram (ECG) signals. Their study demonstrated the feasibility of real-time disease diagnosis using FL at the edge.

4.4 Privacy preservation, security, and ethical needs

Federated healthcare systems must prioritize interoperability, adherence to standards, and implementing robust security measures to ensure seamless collaboration.

4.5 Privacy and security Needs

FL reduces data security and privacy concerns by maintaining local data stores, as opposed to centralized ML techniques, which require datasets to reside on one server (Song et al., 2022). FL is a potential concept for safe, reliable, and impartial models of data. It makes it possible for several parties to work together without exchanging or centralizing datasets (Dasaradharami Reddy & Gadekallu, 2023).

In healthcare, FL involves training ML models on multiple data sources, emphasizing the importance of safeguarding sensitive patient data during processing. It presents several security and privacy challenges; therefore a careful implementation, combined with other privacy-enhancing techniques, is necessary to effectively mitigate the associated risks. Concerns regarding this are discussed below.

4.5.1 Data privacy

FL aims to keep the data localized, restricting its transfer. In FL, instead of transferring data to the central servers, the ML model itself is deployed to each device to be trained on the data (Dasaradharami Reddy & Gadekallu, 2023). However, there is a risk of data leakage, while an adversary could attempt to reconstruct sensitive information by analyzing the model updates during the learning process (Bonawitz, Ivanov, et al., 2019).

4.5.2 Model poisoning attacks

Model poisoning exploits the fact that FL gives malicious participants direct influence over the joint model, enabling much more powerful attacks compared with training-data poisoning. For example, in a healthcare setting, an attacker might introduce biased medical records, leading to harmful decisions and incorrect predictions (Bagdasaryan et al., 2020).

4.5.3 Differential privacy

FL can employ differential privacy techniques to protect individual data privacy. The addition of noise to safeguard privacy may impact the quality and accuracy of predictions, especially in healthcare applications where precision is of utmost importance (Abadi et al., 2016).

4.5.4 Secure model aggregation

In FL, models aggregate the updates from different participants. The aggregation process should be resistant to attacks attempting to extract information from these updates. Safeguarding against collusion attacks poses a significant challenge (McMahan, Moore, et al., 2017).

4.5.5 Data minimization

Another way of minimizing risks is to transfer a minimal amount of data and potentially limit privacy breaches (Yang et al., 2019).

4.5.6 Secure and encrypted communication

Secure communication protocols can be used to transmit data between participants and prevent unauthorized access. Encryption techniques further add in protecting sensitive healthcare data (Bonawitz, Eichner, et al., 2019).

4.6 Ethical needs

FL, with its promising opportunities, is garnering a lot of attention in healthcare. Regulations, such as the Health Insurance Portability and Accountability Act (HIPAA) in the United States or the General Data Protection Regulation (GDPR) in the European Union, ensure the protection of patient rights and privacy.

Nevertheless, several requirements need to be considered to ensure responsible and ethical implementation. This includes the need for unbiased and FAIR (findable, accessible, interoperable, and reusable) data, minimizing the risk of

reidentification of the individuals within the data, and enforcing strict control of what data is used and for what purpose (Voigt & von dem Bussche, 2017).

When complying with GDPR, the federated approach introduces the complexity of identifying and sharing responsibilities with multiple data controllers, conducting data protection impact assessments, and auditing the environments to ensure that the ML models function as expected.

To ensure privacy, techniques such as encryption, anonymization, and pseudonymization must be used, as they help to protect sensitive information during the model training process. It is important to ensure that patients (data subjects) provide explicit consent, and that they are well-informed about the purpose, use, and processing of their data, while retaining the right to opt out at any given time. This brings transparency to data usage and sharing and ensures the incorporation of privacy protection mechanisms (Alysa et al., 2022).

There should be clear guidelines and agreements on data ownership, control, and access to ensure accountability. Robust security measures must be implemented to safeguard data transfer and storage. Policy-based access controls and other cybersecurity protocols must be in place to protect against unauthorized access or data breaches.

In the event of a data breach data processors and controllers must be prepared to minimize the impact and establish planned steps for promptly informing authorities and patients (data subjects).

Establishing ethical guidelines for FL initiatives is crucial. Addressing these ethical needs can help foster trust, protect patient privacy, ensure fairness, and maximize the benefits of FL in healthcare while minimizing the associated risks.

4.7 Role of federated learning in future digital Healthcare 5.0

Healthcare 5.0 is a new era of healthcare that focuses on advanced technologies, personalized care, and patient empowerment. It represents a shift toward a patient-centric approach where individuals are actively involved in managing their own health. The concept leverages cutting-edge technologies to create a seamless healthcare ecosystem that empowers individuals, promotes preventative care, and provides personalized treatment options (patient-centric care) (Rieke et al., 2020).

By training algorithms locally and sharing only model updates rather than raw data, FL addresses data governance challenges while ensuring patient privacy. It enables healthcare organizations to comply with regulations, such as HIPAA or GDPR. FL increases trust and encourages data sharing between healthcare organizations, fostering large-scale collaboration and knowledge sharing in the healthcare ecosystem and enhancing the overall quality of experience/service in healthcare.

One of the key benefits of FL in Healthcare 5.0 is that it enables learning from diverse datasets, enabling knowledge sharing among healthcare organizations with unique patient populations, demographics, and expertise. This allows the creation of robust, generalizable models that adapt to various patient contexts. However, FL is limited by the need for a trusted central server to aggregate model parameters and distribute the global model. To overcome this limitation, researchers are exploring advancements such as secure multiparty computation, cryptographic techniques, and blockchain-based solutions. These aim to improve the security and decentralization of FL in healthcare (Chang et al., 2021).

4.8 Federated learning and blockchain for healthcare

FL and blockchain are two powerful promising technologies that could be combined to revolutionize healthcare. FL enables collaborative model training while maintaining privacy, while blockchain provides a decentralized and secure framework for storing and sharing data. FL combined with blockchain technology offers several benefits in healthcare:

1. FL ensures privacy and security by keeping sensitive patient data localized, minimizing the risk of data breaches. The integration of FL with blockchain further enhances security, providing a tamper-proof and transparent system for data storage and access.
2. The combination enables data integrity and trust through immutable and auditable records on the blockchain, fostering accountability among healthcare stakeholders.
3. Blockchain facilitates interoperability and data sharing, enabling secure FL across institutions and promoting collaboration in healthcare research (Rehman et al., 2022).

Moreover, FL and blockchain empower patients by giving them control over their health data, allowing secure management of permissions and consent. Finally, this integration drives research advancements by aggregating data from multiple sources, creating larger and more diverse datasets that can lead to breakthroughs in disease prediction, treatment development, and precision medicine.

4.9 Federated learning for collaborative robotics in healthcare

FL offers significant potential for collaborative robotics in healthcare, enabling intelligent and adaptive systems while protecting patient privacy. It enables real-time model updates and decentralized patient data sharing, allowing robots to make informed decisions while maintaining confidentiality. Research focuses on

optimizing FL architectures, communication protocols, and privacy-enhancing techniques for collaborative healthcare robotics.

Collaborative robot (cobot) technology has gained significant adoption in the healthcare and medical device industries, serving as a valuable tool for increasing workforce efficiency, streamlining safety procedures, and facilitating improved workflows. Collaborative healthcare robots are automated systems deployed across the medical sector, capable of performing a range of tasks, including administrative tasks, laboratory testing, patient care, and surgical assistance. By bridging the gap caused by labor shortages, these cobots are helping to ease the burden on the medical industry. Demand for automation in healthcare has been driven by the need to reduce the risk of infection to frontline workers and advancements in inpatient care (Dasaradharami Reddy & Gadekallu, 2023).

In particular, cobots in the healthcare sector demonstrate exceptional efficiency in laboratory testing tasks, offering high precision, fast turnaround times, and reduced reliance on manual processes. In addition, cobots are playing a key role in patient care, performing tasks such as medication dispensing, specimen collection, temperature and blood pressure monitoring, and various tests. These capabilities have freed healthcare workers from tedious tasks. They can prioritize urgent matters and effectively optimize their time.

4.10 Federated learning for integration with 6G in healthcare

By 2030, the sixth generation (6G) of mobile technology is expected to be ubiquitous, as it can be integrated into most sectors of the industry, improving the performance of communications standards and enhancing the current communications network infrastructure. Additionally, 6G is expected to leverage more spectrum, providing even lower latency and higher bandwidth transmission capabilities compared with 5G. Also it is anticipated to extend its reach to rural or remote areas currently lacking cellular signals.

Future wireless systems are expected to significantly improve existing wireless capabilities in terms of network throughput, IoT connectivity, latency (from 1 to 10 ms), reliability, availability, energy efficiency, and security.

Moreover, 6G is expected to deliver a 1000-fold improvement in network throughput when compared with 5G technology. This advancement will enable seamless communication and intelligent connectivity for millions of smart devices, as shown in Fig. 4.4. The increased processing power of 6G wireless networks and devices supporting AI will facilitate the proliferation of augmented reality, more advanced imaging and telepresence technology, and more autonomous robots that can communicate with other devices to perform complex tasks.

The technology of 6G can enable high-quality and immersive telemedicine experiences. By integrating FL with 6G, healthcare providers can use distributed

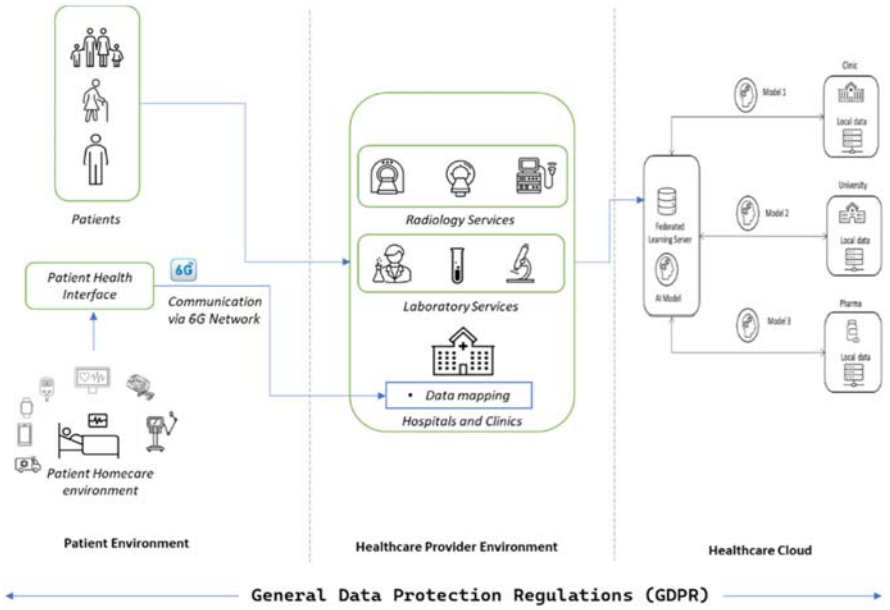

FIGURE 4.4

FL for integration with 6G in healthcare. *FL*, Federated learning.

learning to analyze patient data collected from different remote devices and deliver personalized healthcare services. This combination can support remote monitoring, diagnosis, and treatment, bringing healthcare services closer to patients regardless of their location. It has great potential to revolutionize healthcare delivery, improve patient outcomes, and advance medical research. As 6G technology continues to evolve, further research and innovation will be required to explore the specific applications, challenges, and benefits of integrating FL with 6G in healthcare settings.

4.11 Conclusion

Federated digital health platforms have emerged as an innovative approach to addressing the challenges of privacy and collaboration in healthcare. These systems use the principles of FL, a decentralized ML technique, to enable collaborative analysis and model training on distributed healthcare data.

By enabling collaborative model training, personalized care, and privacy, FL has significant potential for Healthcare 5.0. Its ability to leverage distributed datasets while maintaining security and privacy makes it an attractive approach for advancing healthcare systems toward patient-centric care in the era of Healthcare 5.0.

In summary, federated healthcare platforms utilize distributed data to improve patient privacy and unlock new insights. By utilizing FL, these platforms enable personalized treatments, improved diagnoses, and improved healthcare outcomes. As privacy concerns remain paramount, federated healthcare platforms hold great promise for revolutionizing the healthcare landscape and driving innovation.

Acknowledgment

This work is funded by the European Union under Grant Agreement 101070222. The views and opinions expressed are, however, those of the author(s) only and do not necessarily reflect those of the European Union or the European Commission (granting authority). Neither the European Union nor the granting authority can be held responsible for them.

References

Abadi, M., Chu, A., Goodfellow, I., McMahan, H.B., Mironov, I., Talwar, K., & Zhang, L. (2016). Deep learning with differential privacy. *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security* (pp. 308−318).

Alysa, Z.T., Han, Y., Lizhen, C., & Qiang, Y., Fellow, IEEE. (2022). *Towards personalized federated learning IEEE transactions on neural networks and learning systems*.

Bagdasaryan, E., Veit, A., Hua, Y., Estrin, D., & Shmatikov, V. (2020). How to backdoor federated learning. In *Proceedings of the 1st conference on machine learning and systems* (pp. 265−278).

Bonawitz, K., Eichner, H., Grieskamp, W., Huba, D., Ingerman, A., Ivanov, V., & Wang, S. (2019). Towards federated learning at scale: System design. *arXiv preprint arXiv*, *1902*, 01046.

Bonawitz, K., Ivanov, V., Kreuter, B., Marcedone, A., McMahan, H.B., Patel, & Zhang, Z. (2019). Practical secure aggregation for privacy-preserving machine learning. In *Proceedings of the 2017 ACM SIGSAC conference on computer and communications security* (pp. 1175−1191).

Chang, Y., Fang, C., & Sun, W. (2021). A blockchain-based federated learning method for smart healthcare. *Computational Intelligence and Neuroscience*, *2021*.

Dasaradharami Reddy, K., & Gadekallu, T. R. (2023). A comprehensive survey on federated learning techniques for healthcare informatics. *Computational Intelligence and Neuroscience*, *2023*.

Huang, L., Shea, A. L., Qian, H., Masurkar, A., Deng, H., & Liu, D. (2019). Patient clustering improves efficiency of federated machine learning to predict mortality and hospital stay time using distributed electronic medical records. *Journal of Biomedical Informatics*, *99*, 103291.

Li, T., Sahu, A. K., Talwalkar, A., & Smith, V. (2020). Federated learning: Challenges, methods, and future directions. *IEEE Signal Processing Magazine*, *37*(3), 50−60.

Li, X., Wang, Z., Wu, T., Jiang, H., & Pang, C. (2020). Privacy-preserving chest X-ray classification with split learning and federated learning. *IEEE Access*, *8*, 147937−147945.

McMahan, B., Moore, E., Ramage, D., Hampson, S., & y Arcas, B. A. (2017). *Communication-efficient learning of deep networks from decentralized data*. *Artificial intelligence and statistics* (pp. 1273–1282). PMLR.

McMahan, H.B., Ramage, D., Talwar, K., & Zhang, L. (2017). Learning differentially private recurrent language models. In *Proceedings of the 2017 conference on empirical methods in natural language processing* (pp. 2792–2802).

Nguyen, D. C., Ding, M., Pathirana, P. N., Seneviratne, A., Li, J., & Poor, H. V. (2021). Federated learning for internet of things: A comprehensive survey. *IEEE Communications Surveys & Tutorials*, *23*(3), 1622–1658.

Rehman, A., Abbas, S., Khan, M. A., Ghazal, T. M., Adnan, K. M., & Mosavi, A. (2022). A secure healthcare 5.0 system based on blockchain technology entangled with federated learning technique. *Computers in Biology and Medicine*, *150*, 106019.

Rehman, Z. U., Zia, M. S., Bojja, G. R., Yaqub, M., Jinchao, F., & Arshid, K. (2020a). Texture based localization of a brain tumor from MR-images by using a machine learning approach. *Medical Hypotheses*, 109705.

Rieke, N., Hancox, J., Li, W., Milletari, F., Roth, H. R., Albarqouni, S., & Cardoso, M. J. (2020). The future of digital health with federated learning. *NPJ Digital Medicine*, *3*(1), 119.

Shen, D., Wu, G., & Suk, H.-I. (2017). Deep learning in medical image analysis. *Annual Review of Biomedical Engineering*, *19*, 221–248.

Silva, S., Gutman, B. A., Romero, E., Thompson, P. M., Altmann, A., & Lorenzi, M. (2019). *Federated learning in distributed medical databases: Meta-analysis of large-scale subcortical brain data*. *2019 IEEE 16th international symposium on biomedical imaging (ISBI 2019)* (pp. 270–274). IEEE.

Song, J., Wang, W., Gadekallu, T. R., Cao, J., & Liu, Y. (2022). EPPDA: An efficient privacy-preserving data aggregation federated learning scheme. *IEEE Transactions on Network Science and Engineering*.

Song, J., Wang, W., Gadekallu, T. R., Cao, J., & Liu, Y. (2022). Eppda: An efficient privacy-preserving data aggregation federated learning scheme. *IEEE Transactions on Network Science and Engineering*.

Voigt, P., & von dem Bussche, A. (2017). *The EU General Data Protection Regulation (GDPR)*. Springer International Publishing.

Yang, Q., Liu, Y., Chen, T., & Tong, Y. (2019). Federated machine learning: Concept and applications. *ACM Transactions on Intelligent Systems and Technology (TIST)*, *10*(2), 1–19.

Zhang, W., Zhou, T., Lu, Q., Wang, X., Zhu, C., Sun, H., & Wang, F. Y. (2021). Dynamic-fusion-based federated learning for COVID-19 detection. *IEEE Internet of Things Journal*, *8*(21), 15884–15891. Available from https://www.theodi.org/article/federated-learning-an-introduction-report/.

This page intentionally left blank