CAN FEDERATED LEARNING SOLVE AI'S DATA PRIVACY PROBLEM?: A LEGAL ANALYSIS

52 Rutgers L. Rec. 252 (2025) | WestLaw | LexisNexis | PDF

The speed of development in Artificial Intelligence (AI) in recent years has been breathtaking. Yet this comes with its own set of problems.[1 One of these problems is that the current iteration of AI requires vast quantities of data for the training of AI systems, and, currently, the demand for data is outstripping the supply of data.[2 This may hinder the further development and improvement of AI systems. This is sometimes referred to as AI's data problem.[3 However, making more data available to train AI raises various concerns, not least, in relation to data privacy, thus, AI has a data privacy problem.

Federated learning (FL) may provide a solution to this problem.[4 The basic premise behind FL is simple: In the standard training of AI systems, data is collected and transferred onto a central server, where the AI system trains on the data. In contrast, in FL, data is not collected but remains in its original locations. Instead, each party receives the raw model, which is then trained on the dataset in situ. Upon completion of the training, the trained model is sent back, and is combined with other similarly trained models, into a single, integrated model. The result is that the integrated model has effectively been trained on all the datasets, but no data is transferred out from its original location.

Data is often held in ?silos'.[5 A data silo is anything that holds data (e.g. smartphones, laptops, hospitals, banks, etc.) but accessing data held in silos is challenging. Sometimes regulation, like data protection legislation or IP law, prevents data from being shared. Alternatively, there may be a reluctance to share data, for instance, due to concerns about data confidentiality or data integrity. The intended purpose of FL is to ?break open' these data silos, by enabling the training of AI systems while preserving data privacy and confidentiality. If FL can fulfill this promise, this could bring significant benefits. By way of example, ?healthcare providers could train algorithms to develop new drugs based on patient data, while maintaining privacy and patient confidentiality, or researchers in different countries could train algorithms without transmitting data across jurisdictions.?[6

FL has generated significant interest amongst the computer science community, however, there is a dearth of writings on, and understanding of, FL among lawyers and legal academics.[7 This is a problem because the conceptualization of concepts like data and privacy may differ across disciplines[8 and to what extent FL can break open data silos created by regulation requires a legal analysis. This article aims to fill this lacuna by providing a comprehensive legal analysis of FL. This will be done by examining how the data protection principles ? represented by the most stringent standards under the European Union's General Data Protection Regulation (GDPR)[9 ? applies to FL. The argument will made that from a legal perspective FL can indeed be an effective method to ensure compliance with data protection regulation.

Although the legal analysis in this article focuses on the GDPR, the significance of the analysis extends beyond the EU. EU regulation has proven influential beyond the EU,[10 and many data protection regimes are modelled on the GDPR.[11 Moreover, FL raises an important conceptual question about the relationship between data protection and the development of AI; that is, whether the training of AI systems on personal data is in itself an infringement of data protection rights, or whether there is such an infringement only because of some feature of how the training is conducted, e.g. that data is collected to a central server or access to the data is given to a third party. In the standard training of AI systems, this question will seldom arise as data needs to be collected for the AI systems to be trained, and many data protection regimes regulate the collection of data.[12 However, because in FL no data is collected, this issue is brought into sharp focus, and in the age of AI, this is a question every data protection regime will need to answer.[13

This article suggests that the training of AI systems itself does not infringe data protection rights, provided that the data is kept secure from abuse (i.e. the data being used for purpose other than training AI). The argument is that using personal data to train AI systems does not reveal information about an individual, such information is only revealed when the AI system is applied to a particular case. This article will show that the GDPR can be interpreted in this way, and if this interpretation is followed, the GDPR can provide for the protection of personal data, without hindering the development of AI systems.[14 For the legal analysis of FL

this means that the question of whether the training of an AI system through FL is GDPR compliant will largely depend only on one factor, namely whether the data is kept secure, rather than the host of factors, which is typically required to assess GDPR compliance in standard training of AI systems. Thus, FL should make it easier for AI developers to train models on personal data.

Despite FL being a potential boon to AI's compliance with data protection regulation, this article will express doubt as to whether FL can make a significant contribution towards solving AI's data problem. Although, FL may be an effective way to deal with data protection, data protection is only one among other obstacles to data sharing. For instance, IP law may prevent data from being shared, and FL does not directly impact the application of IP law. Moreover, it is unlikely that FL will be used sufficiently widely to make significantly more data available, than is currently the case. There is also a lack of legal clarity in relation to FL, and without legal clarity it is unlikely that FL will be commonly adopted. Furthermore, currently, FL is not used widely across different organizations.[15 This means that a lot of data will remain inaccessible. This is an area where regulators and policy makers may be able to make a positive contribution. This paper will suggest that, if regulators and policy makers decide to facilitate the use of FL, a possible tool is the creation of a FL regulatory regime, including an FL licensing regime, to facilitate data sharing across organizations.

This article will proceed as follows. First, an overview of FL will be provided. Second, this article will give a detailed analysis of how the GDPR applies to the training of AI systems. This analysis will take up considerable space, but it is crucial to understand how the GPPR applies to the training of AI systems as without such understanding, it is impossible to examine how the GDPR applies to FL. Third, this article will analyze to what extent FL can facilitate the sharing of non-personal data and examine the obstacles to FL being used more widely. Note that this article focuses on FL in relation to data protection regulation. Other issues, like IP law or antitrust law, will not be considered directly and are out of the scope for this article.

[1 Thilo Hagendorff & Katharina Wezel, 15 Challenges for AI: Or What AI (Currently) Can't Co, 35 AI & Soc 355 (2020).

[2 Tal Roded & Peter Slattery, What Drives Progress in AI? Trends in Data, FutureTech (March 19, 2024), https://futuretech.mit.edu/news/what-drives-progress-in-ai-trends-in-data.

[3 Devika Rao, All-powerful, ever-pervasive AI is running out of internet, The Week (June 5, 2024), https://theweek.com/tech/ai-running-out-of-data; S.E. Whang, et al. Data collection and quality challenges in deep learning: a data-centric AI perspective, 32 VLDB J. 79 (2023).

[4 See Brendan McMahan & Daniel Ramage,, Federated Learning: Collaborative Machine Learning without Centralized Training Data, Google Research Blog (April 6, 2017), https://ai.googleblog.com/2017/04/federated-learning-collaborative.html.

[5 See Florian Gamper, Federated Learning: What Lawyers Need to Know, L.Gazette, (June 2024), https://lawgazette.com.sg/feature/federated-learning-what-lawyers-need-to-know/ (the Law Gazette is the official publication of the Law Sciety of Singapore).

[6 Id.

[7 However, there is some legal analysis of FL. See e.g. S. Rossello et al., Data Protection by Design in AI? The Case of Federated Learning, 116 Computerrecht (2021); Nguyen Truong et al. Privacy Preservation in Federated Learning: An Insightful Survey from the GDPR perspective, 110 Comput. Secur. J. 12402, 12414-18 (2021).

[8 The same claim could be made in relation to many other concepts, like transparency, bias, fairness, to mention just a few.

[9 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) 2016 O.J. (L119) [hereinafter GDPR.

[10 Anu Bradford, THE BRUSSELS EFFECT, 107 Nw. U. L. Rev. 1 (2012) (argues that EU regulation impact jurisdictions outside the EU).

[11 Graham Greenleaf, Now 157 Countries: Twelve Data Privacy Laws in 2021/22, 176 Privacy L. & Bus. Int'l Rep., 1, 1 (2022).
[12 GDPR, supra note 11, art. 4.1(2), (states that collecting is a form of processing, inter alia GDPR arts. 5 and 6 regulate processing).

[13 Just to clarify, the question is also relevant for jurisdictions which currently do not have a data protection regime but are considering creating such a regime.

[14 See Giovanni Sartor & Francesca Lagioia, The impact of the General Data Protection Regulation (GDPR) on artificial intelligence, European Parliamentary Rsch. Serv, 76 (June 2020) [hereinafter EPRS Study (A study at the request of the Panel for the Future of Science and Technology (STOA) and managed by the Scientific Foresight Unit, within the Directorate-General for Parliamentary Research Services (EPRS) of the Secretariat of the European Parliament).

[15] Saikishore Kalloori & Abhishek Srivastava, Towards cross-silo federated learning for corporate organizations, 289 Knowledge-Based Sys. 1, (Apr. 8, 2024).

View the Entire Article