POSITION PAPER

Artificial Intelligence for the public sector: Masked Federated Learning as a new privacy-protecting solution

Martin Schallbruch
(Digital Society Institute at the ESMT Berlin)

Professor Michael Huth
(Imperial College London, Xayn AG)

Dr. Leif-Nissen Lundbæk
(Xayn AG)

Dr. Clara Herdeanu
(Xayn AG)

Lola Attenberger
(Digital Society Institute at the ESMT Berlin)

TABLE OF CONTENTS

3 . . . . **1. INTRODUCTION**

4 . . . . **2. MASKED FEDERATED LEARNING:** Decentralised technology with built-in privacy protection

6 . . . . **3. POTENTIAL USE CASES** from the public sector

3.1. Public health: After the pandemic is before the pandemic

3.2. Law enforcement: Early detection of crimes through intelligent data analysis

10 . . . . **4. OUTLOOK**

11 . . . . **5. SOURCES**

Imprint

1. INTRODUCTION

If we want to future-proof the state and its administration and make them more efficient, we need technology. The Coronavirus pandemic has shown that the digitalisation of public authority structures in Germany still falls far short of expectations. In the coming years, it will therefore be important to consistently digitalise the public authorities and administration at the federal, state, and local level.

Artificial intelligence (AI) holds great potential for this digitalisation effort. Machine learning and AI can contribute significantly to optimising and automating decision-making processes in administrations, organisations, and institutions, so that they become faster and more cost-effective. With the help of AI, public administrations can provide their internal and external services in a more targeted, tailored and simple manner[1].

But any form of digitalisation of public services and digital cooperation between authorities, companies, and citizens must meet **high standards of data protection and trustworthiness**. Unnecessary central data collection, excessive storage and use of personal data, or non-transparent data use are incompatible with people's expectations of data protection and self-determination. In addition, this also contradicts European data protection law. To use AI meaningfully in the public sector and to create high benefits for the community, the innovative potential of AI must be combined with data economy, transparency, and trustworthiness.

New technical approaches are available that meet these criteria. In the following, the authors outline one of the most promising approaches, known as **Masked Federated Learning**, and explain its application in the public sector using the examples of healthcare and law enforcement.

[1] *Cf. Die Bundesregierung (2018): Strategie Künstliche Intelligenz der Bundesregierung. Stand: November 2018. p. 31. Link*

2. MASKED FEDERATED LEARNING: Decentralised technology with built-in privacy protection

To train AI models, large amounts of data are needed – in many cases also personal data. The more such personal data is used, the greater the privacy risks. This tension leads to the **"AI privacy dilemma"**, which can be explained as follows: Collecting more data leads to more powerful and precise AI applications but also poses greater privacy risks. Reducing the amount of data promotes privacy protection but also results in inferior technology that does not use the full potential of AI-based applications.

One solution to the dilemma is **Masked Federated Learning**[2]**, a decentralised technology that combines data protection with powerful AI**. With this approach, the raw data remains where it is generated – on users' end devices. Instead of bringing the data to the algorithms, the algorithms are sent to the data and trained on-site. These AI models trained on the local data sets are then aggregated into a global model. This global model then serves as the basis for subsequent local learning by feeding back the collective experience to the local models, so that a form of collaborative learning takes place (cf. Figure 1[3]). Only encrypted AI models are communicated in the network, without any sensitive personal data. Since the raw data never leaves the end devices, the costly and technically dubious anonymisation of the data becomes superfluous.

Masked Federated Learning, therefore, meets the **requirements of data protection directives such as the GDPR**. In addition, the approach offers the advantage that learning can take place asynchronously since the individual local models are relatively independent from each other. Due to the exponential growth of global data streams, it is not sustainable to store all this information centrally for further processing. Asynchronous decentralised learning scales far better

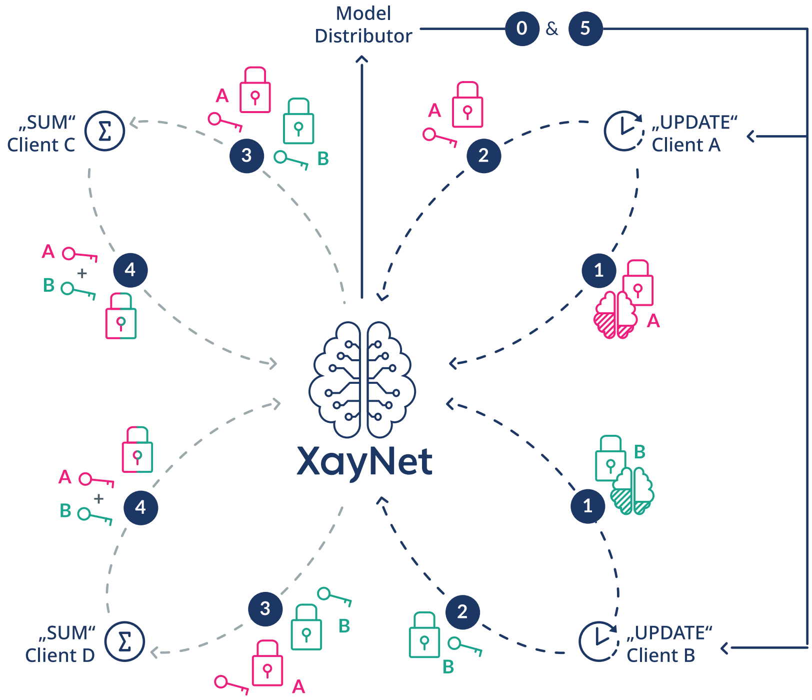

Fig. 1: Masked Federated Learning (Source: Xayn AG)

[2] Cf. Dänschel, L.; Huth, M.; Lundbæk, L. (2020): XayNet: Masked Cross-Device Federated Learning Framework. Link
[3] Ibid. p. 16.

with growing volumes of data. The approach also offers advantages for IT security: potential attackers would have to target many different end devices simultaneously to be successful.

Masked Federated machine learning can thus condense the growing data streams into suitable AI models, **protect personal data, ensure the performance and resilience of the AI, and be highly scalable at the same time**.

Masked Federated Learning can be applied for "cross-silo" use cases, i.e., cross-silo solutions in which a smaller number of data sets are hosted in powerful environments such as servers, as well as "cross-device" use cases in which cross-device solutions with many millions of individual devices are implemented. The stored data sets can vary significantly in size and patterns. For cross-device applications in particular, it might be problematic to switch from a centralised to a decentralised AI model. In addition, cross-device approaches also require a minimum performance of the end devices.

In the following, the authors illustrate the potential of this method with examples from the health sector and law enforcement, highlighting the **particularly high demands on data protection and resilience for the public sector**.

3. POTENTIAL USE CASES from the public sector

3.1. Public health: After the pandemic is before the pandemic

The Coronavirus pandemic has turned the world upside down and has already claimed more than 3 million lives globally[4]. International experts are already warning that we have entered an era of pandemics and that they will occur more frequently in the future and be faster and more deadly[5]. The Covid-19 pandemic must thus be seen as a warning shot for future pandemics. Decision-makers should prepare themselves for further serious cases. In Germany, the fight against Covid-19 has given digitalisation an enormous boost, but at the same time it has also revealed digital weaknesses in our country – such as the use of new technologies for public health.

The "Corona Warn App"[6] (Coronavirus warning app) has shown that centralised technologies can help protect individual and public health in a privacy-preserving way. Such approaches should be developed further. It is conceivable that AI could be used to go beyond Covid-19 and provide individuals with a tool to identify their specific risk for infectious diseases. A digital application based on Masked Federated Learning can provide personal assessments based on the respective preexisting conditions such as physical constitution, sex, age, regional infection history, existing contacts with (potentially) infected persons, and vaccination status. This highly sensitive personal data remains on users' end devices, while only the encrypted AI models are used for a continuous improvement of the AI's assessment.

In addition, assessments are bundled in this way to reach conclusions on potential infection sources and risks at a systemic level. The health authorities could use such collective results to better assess which measures provide the best health protection. For this purpose, the health departments don't need and don't receive the highly sensitive personal raw data.

If during a future pandemic new scientific findings on risk factors emerge, they could be fed into the model, and the people affected by exogenous and endogenous factors could be warned individually. **Such a decentralised approach in combination with cross-device and cross-silo applications would offer better individual protection to individuals. At the same time, health departments and other authorities would be empowered to make evidence-based decisions to better protect the population.**

The more individuals install and use such an application, the greater the improvement potential for decisions by health departments and policymakers will be.

3.2. Law enforcement: Early detection of crimes through intelligent data analysis

In a largely digitalised world, the analysis of large amounts of digital data is a necessary and effective aid to police work. Anomalies in data files can help detect crimes at an early stage, confirm existing suspicions, or enable law enforcement to take new investigative approaches. Police can draw on various data sources for this purpose.

Within the framework of the project Police 2020, a 'data warehouse' of the German

[4] *World Health Organization (2021): WHO Coronavirus (Covid-19) Dashboard. Link*

[5] *Cf. The Intergovernmental Science Policy Platform on Biodiversity and Ecosystem Services (2020): IPBES Workshop Report on Biodiversity and Pandemics. United Nations Environment Programme. Link*

[6] *Cf. Die Bundesregierung (2020b): Corona Warn App – Coronavirus warning app. Link*

police is to be created (cf. Figure 2[7]), in which a common nationwide data inventory (cf. Figure 2, green) will be supplemented by data inventories of the individual police authorities. In the future, the common database will (be able to) grow, but not all the knowledge of the police authorities will be available to all other authorities.

When police officers are working on a case, they collect a lot of digital data on the spot – including, for example, data from searches in social media, from seized hardware, from radio cell searches or telecommunications surveillance. In addition, law enforcement agencies might be able to access existing state data, such as citizen or vehicle registers. Finally, there are several legally standardised suspicious activity reports that oblige companies to directly or indirectly share digital data with law enforcement authorities, so that they can detect certain criminal offences at an early stage. This includes suspicious activity reports from banks for money-laundering prevention or suspicious activity reports on hate crimes from social media platform operators. In addition, internet providers will be obliged to report certain cybercrimes to the Federal Criminal Police Office.

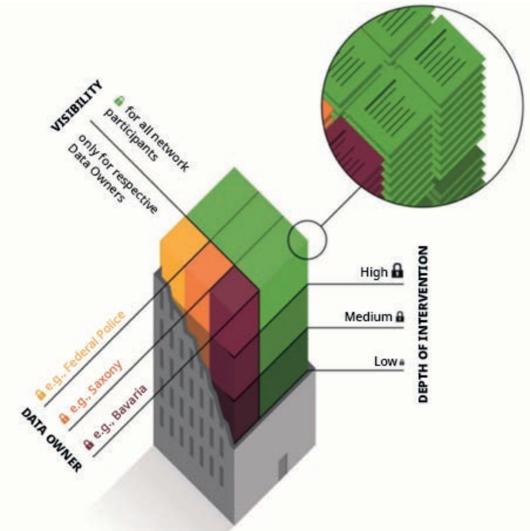

Fig. 2: Data warehouse of the German police (Source: Federal Ministry of the Interior[8])

Police authorities face the task of analysing and evaluating the incoming information, from which indications of potential criminal offences arise. **Given the growing abundance and diversity of information, this evaluation requires a high degree of automation.** In the first step, information must be classified as automatically as possible to determine whether it appears to be relevant to a certain criminal activity. Only if this is the case should the information be checked manually and investigations carried out. Such an initial classification of information differs according to the area of the criminal activity and the type of information.

In the case of offences involving images of sexual violence against children and adolescents[9], the first step is to check which images can provide evidence of a crime. The first step in examining money laundering is to single out specific data records from a set of payment transaction data that are relevant for further analysis. To identify fraudulent e-commerce activities, those that correspond to certain patterns of fraudulent transactions must be filtered out of a set of secured communications data.

The quality of such initial, largely automatic analyses of digital data depends highly on whether and to what extent it is possible to leverage police experience in the respective offences, their specific patterns and accompanying circumstances. All offences mentioned as examples are characterised by certain modes of operation, which can be used to classify the data. What kind of representations are common in depictions of sexual violence

[7] *Bundesministerium des Innern: Polizei 2020. White Paper, Januar 2018. Link*

[8] *Bundesministerium des Innern: Polizei 2020. White Paper, Januar 2018, p. 12. Link*

[9] *The authors deliberately do not use the common wording "child pornography", as this term trivialises the underlying sexualised violence against children and adolescents – see also Frauen gegen Gewalt e.V.: Notes for reporting on violence against women and children. Link*

against children and adolescents? Which transactions have a high potential for money laundering? What patterns of enquiry are characteristic of bots that want to defraud online traders? These kinds of police experience values solidify with the number of cases that are processed. At the same time, they must be open to dynamic development and adapt to new forms of crime quickly.

AI can help analyse and classify such data – as already shown by various projects at German security authorities[10]. The ongoing projects are within the remit of a single authority and relate to a specific criminal phenomenon. This rather selective approach does not utilise the full potential that lies in the accumulated experiential knowledge of all police authorities. It would therefore make sense to have a solution that offers AI support in the classification of information in a way in which the experience of all authorities can be leveraged collectively. However, only in the case of particularly serious crimes and with explicit legal authorisation is it possible to bring together all phenomenal data from all police authorities to develop comprehensive AI models.

The federated learning approach offers a solution **not to exchange personal information between agencies but only the experiential knowledge in the classification of information**. In this way, a police department benefits from the experience of other departments. This increases the experience and knowledge of the police force as a whole without the central collection of all personal data. The decentralised models jointly train a centrally operated model that is available to all police departments. The global model is then used to improve the local models **– all without the police forces having to share the raw personal data**.

[10] Current AI projects of German security authorities are, for example, the automated image recognition in investigations of sexualised violence against children and adolescents (StA Köln, cf. Behördenspiegel, January 2021, p. 29) or the AI-based system "FIU Analytics" of the Financial Intelligence Unit (FIU) for money laundering detection (cf. Handelsblatt, 24 November 2020).

4. OUTLOOK

Some consider the European Union probably the most trustworthy region in the world because of its data protection regulation and see this as a competitive advantage[11]. Others regard data protection as hampering innovation[12] or even damaging in times of a pandemic[13]. The afore explained **Masked Federated Learning is ideally suited to reconcile the protection and the use of personal data**. As a decentralised self-learning technique, this distributed learning follows a **privacy-by-design** approach, since no personal data needs to be centrally aggregated. The data is only stored locally on the user's device.

This approach can support the implementation of the official German AI strategy[14], which on the one hand, promotes further applications, especially in the public sector, and on the other hand attaches particular importance to the design of AI applications in conformity with fundamental rights. The German government aims to use data innovatively without jeopardising data protection[15].

For healthcare and nursing, the government is not only concerned with research projects, but with data-supported, AI-assisted healthcare applications in everyday life. The given potential use case for the public health sector shows **how this technology could be used to quickly identify and contain risks in future pandemics and ensure a faster response capability**. The collaborative, decentralised technology could support citizens and health departments in the exchange of information and data analysis. The system could enable citizens to assess risks faster and in a better-informed way. At the same time, communication between the authorities and with the service users could be significantly more effective. For public security, the Federal Government considers AI to be a strategic instrument to further improve the performance of the German police, for example in the case of hybrid threats. The examples shown illustrate that **distributed federated learning could facilitate police work as a whole without centrally pooling data**. The fundamentally problematic disclosure of personal data in the run-up to concrete suspicions could be avoided, but machine learning could still be used to support police investigations.

Overall, the use cases for the public health sector and the police demonstrate the potential that a de-centrally organised self-learning system offers for public administration. Administrative services for citizens could be improved through distributed learning, as **the technology has the potential to increase efficiency and strengthen citizens' trust in the administration**. Masked Federated Learning could thus become an **EU model for responsible AI use**. All the advantages of machine learning can be exploited – while at the same time **personal data is protected**.

[11] *Cf. Dreo, G.; Eiseler, V.; Gentschen Felde, N.; Gehrke, W.; Helmbrecht, U.; Hommel, W.; Zahn, J. (2020): Europäische Digitale Souveränität: Weg zum Erfolg? – Ein Bericht zur Jahrestagung CODE 2020. In: Z Außen Sicherheitspolitik 13: 399-404. Link*

[12] *Cf. Mertens, P. (2019): Die Datenschutz-Grundverordnung – eine kritische Sicht. In: Wirtschaftsinformatik & Management, 11(1): 6-17. Link*

[13] *Cf. Hottelet, U. (2021): Corona-Pandemie: Bremst uns der Datenschutz aus? Link*

[14] *Cf. Die Bundesregierung (2020a): Strategie Künstliche Intelligenz der Bundesregierung. Fortschreibung 2020. Stand: Dezember 2020. p. 22. Link*

[15] *Ibid.*

5. SOURCES

- Dänschel, L.; Huth, M.; Lundbæk, L. (2020): XayNet: Masked Cross-Device Federated Learning Framework. Online: https://uploads-ssl.webflow.com/5ea197660b956f76d26f0026/5fcf7f97c0333cb84277fcd8_XayNet%20Whitepaper%203.0%20v3.pdf

- Bundesministerium des Inneren (2018): Polizei 2020. White Paper, Januar 2018. Online: https://www.bmi.bund.de/SharedDocs/downloads/DE/veroeffentlichungen/2018/polizei-2020-white-paper.html

- Die Bundesregierung (2018): Strategie Künstliche Intelligenz der Bundesregierung. Stand: November 2018. Online: https://www.bmwi.de/Redaktion/DE/Publikationen/Technologie/strategie-kuenstliche-intelligenz-der-bundesregierung.pdf?__blob=publicationFile&v=10

- Die Bundesregierung (2020): Strategie Künstliche Intelligenz der Bundesregierung. Fortschreibung 2020. Stand: Dezember 2020. Online: https://www.bmwi.de/Redaktion/DE/Publikationen/Technologie/strategie-kuenstliche-intelligenz-fortschreibung-2020.pdf?__blob=publicationFile&v=10

- Dreo, G.; Eiseler, V.; Gentschen Felde, N.; Gehrke, W.; Helmbrecht, U.; Hommel, W.; Zahn, J. (2020): Europäische Digitale Souveränität: Weg zum Erfolg? – Ein Bericht zur Jahrestagung CODE 2020. In: Z Außen Sicherheitspolitik 13: 399-404. Online: https://link.springer.com/article/10.1007%2Fs12399-020-00829-2

- Die Bundesregierung (2020b): Corona Warn App – Coronavirus warning app. Online: https://www.bundesregierung.de/breg-de/themen/corona-warn-app/corona-warn-app-englisch

- Feldmann, M; Proll, Uwe (2021): Kein Ersatz für den Menschen. In: Behördenspiegel, Januar 2021, S. 29. Online: https://issuu.com/behoerden_spiegel/docs/2021_januar

- Frauen gegen Gewalt e.V.: Hinweise für die Berichterstattung über Gewalt gegen Frauen und Kinder. Online: https://www.frauen-gegen-gewalt.de/de/ueber-uns/presse/informationen-fuer-die-presse/hinweise-fuer-die-berichterstattung-ueber-gewalt-gegen-frauen-und-kinder.html

- Hottelet, U. (2021): Corona-Pandemie: Bremst uns der Datenschutz aus? Online: https://www.heise.de/news/Corona-Pandemie-Bremst-uns-der-Datenschutz-aus-5042875.html

- Mertens, P. (2019): Die Datenschutz-Grundverordnung – eine kritische Sicht. In: Wirtschaftsinformatik & Management, 11(1): 6-17. Online: https://link.springer.com/article/10.1365/s35764-019-0159-5

- Stiens, T. (2020): Wie Software künftig bei der Geldwäsche-Bekämpfung helfen soll. In: Handelsblatt 24.November 2020. Online: https://www.handelsblatt.com/politik/deutschland/finanzkriminalitaet-wie-software-kuenftig-bei-der-geldwaesche-bekaempfung-helfen-soll/26596064.html?ticket=ST-5862866-XOIrpKJFJNMLfhA1nftg-ap3

- The Intergovernmental Science Policy Platform on Biodiversity and Ecosystem Services (2020): IPBES Workshop Report on Biodiversity and Pandemics. United Nations Environment Programme. Online: https://www.unep.org/resources/report/ipbes-workshop-report-biodiversity-and-pandemics

- World Health Organization (2021): WHO Coronavirus (Covid-19) Dashboard. Online: https://covid19.who.int/

ABOUT XAYN
www.xayn.com

Xayn is a privacy-protecting search engine that enables users to gain back control over the algorithms and provides them with a smooth user-experience. Using the latest AI technology made in Europe, the company ushers in a new generation of user-friendly privacy tech – making privacy available for everyone.

The AI company started as a research project at The University of Oxford and Imperial College London by Leif-Nissen Lundbæk (Ph.D.) and Professor Michael Huth. Together with Felix Hahmann, they founded the tech company in 2017. To this day, that academic vision remains with a workforce comprised of 30% Ph.D.s The company's open-source framework for federated analytics and learning, XayNet, is the basis of the privacy-protecting personalised search alternative Xayn. The Berlin-based company has received investment funding of 9.5 million EURO by Earlybird VC as well as Dominik Schiener. Xayn has worked with corporations such as Porsche, Daimler, Deutsche Bahn, and Siemens.

ABOUT ESMT BERLIN
esmt.berlin

ESMT Berlin is the highest-ranked business school in Germany and Top 10 in Europe. Founded by 25 leading global companies, ESMT offers master's, MBA, and PhD programs, as well as executive education on its campus in Berlin, in locations around the world, online, and in online blended format. Focusing on leadership, innovation, and analytics, its diverse faculty publishes outstanding research in top academic journals. Additionally, the international business school provides an interdisciplinary platform for discourse between politics, business, and academia. ESMT is a non-profit private institution of higher education, accredited by AACSB, AMBA, EQUIS, and FIBAA, and is committed to diversity, equity, and inclusion across all its activities and communities.

ABOUT THE DIGITAL SOCIETY INSTITUTE

The Digital Society Institute (DSI) is an interdisciplinary research institute of ESMT Berlin, founded in 2015 with the support of leading global companies. It accompanies the economic and social design of digitalization through strategic research and development – with a strong focus on cybersecurity.

IMPRINT

Publisher:
Xayn AG
Unter den Linden 42
10117 Berlin
press@xayn.com

Editor: Dr. Clara Herdeanu

Layout: Julia Hintz

All internet links mentioned in the text were checked on 04 June 2021.

Berlin, June 2021