6-18-2022

TRACING DOWN THE VALUE OF CO-CREATION IN FEDERATED AI ECOSYSTEMS

Marco Röder
University of Wuerzburg, marco.roeder@uni-wuerzburg.de

Peter Kowalczyk
University of Wuerzburg, peter.kowalczyk@uni-wuerzburg.de

Frédéric Thiesse
University of Wuerzburg, frederic.thiesse@uni-wuerzburg.de

TRACING DOWN THE VALUE OF CO-CREATION IN FEDERATED AI ECOSYSTEMS

Research Paper

Röder, Marco, University of Würzburg, Würzburg, Germany, marco.roeder@uni-wuerzburg.de

Kowalczyk, Peter, University of Würzburg, Würzburg, Germany, peter.kowalczyk@uni-wuerzburg.de

Thiesse, Frédéric, University of Würzburg, Würzburg, Germany, frederic.thiesse@uni-wuerzburg.de

Abstract

Methods and software components for developing novel IT solutions based on artificial intelligence (AI) technology are broadly available to organizations of any size. However, as AI typically requires large amounts of data, smaller organizations are at a disadvantage compared to large competitors as the latter often have more training and test data at their disposal. Collaboration and data sharing between multiple smaller actors might offer a solution to this issue, but also poses a potential risk to privacy and confidentiality. Our research considers the concept of federated learning, which enables collaborative training without exchanging the actual data. Still, the benefits of value co-creation within federated AI ecosystems are unclear. To shed light on this topic, we present a data-driven analysis using the example of sales forecasting in retail. We show that three types of benefits can be expected in federated AI ecosystems, namely collaboration, privacy preservation, and generalizability.

Keywords: Federated Learning, Value Co-creation, AI Ecosystem, Deep Learning.

1 Introduction

Artificial intelligence (AI) has a steadily increasing influence on our everyday life (Schmidt et al., 2019). This holds, among others, for the retail sector (Grewal et al., 2021; Guha et al., 2021), where consumer products can nowadays easily be ordered by talking to a voice assistant, smart recommender systems suggest suitable complementary products, and chat-bots provide adequate answers to any detailed question about a specific product type. The proliferation of AI in the retail industry is expected to continue in the years to come due to the rich potential for both new market entrants and well-established companies alike to shape and boost their competitive position (Grewal et al., 2017; Guha et al., 2021; Jöhnk et al., 2021). A recent report from McKinsey & Company (Chui et al., 2018) indicates that the use of the more sophisticated deep learning (DL) technologies as opposed to rather traditional AI technologies will further increase the added value by an estimated 87%.

However, despite this economic potential, some retailers shy away from using advanced DL technologies (Mahmoud et al., 2020; Oosthuizen et al., 2021; Shankar et al., 2021), which may be attributed to multiple reasons. Besides the sheer complexity and lack of immediate interpretability of the generated models (Adadi and Berrada, 2018; Oosthuizen et al., 2021; Samek and Müller, 2019), DL also requires large amounts of data to achieve high-level performance if trained from scratch (Bengio et al., 2017; Najafabadi et al., 2014). This may become particularly difficult for small and medium-sized enterprises (SMEs) which, regarding the availability of sufficient heterogeneous data, are often at a disadvantage compared to their larger competitors when it comes to training a data-hungry DL model (Bauer et al., 2020; Ulrich et al., 2021). The issue of insufficient data for training and testing may ultimately lead to a competitive disadvantage for retailers that do not have rich and accessible data pools at their disposal or even turn into a showstopper for new entrants. Against this backdrop, the collaboration of multiple companies through data sharing seems like a promising remedy. However, as data sharing may pose a severe data fraud risk and violate data privacy rights, many companies are reluctant to collaborate with others in their DL projects (Mahmoud et al., 2020; Sheller et al., 2020).

The emerging concept of federated learning (FL) promises to mitigate these risks (McMahan et al., 2017). FL denotes a new machine learning (ML) approach where a central service provider connects and coordinates multiple data owners (e.g., companies) in the joint training of a large DL model through the exchange and aggregation of just the estimated model weights but not the underlying raw data (Kairouz et al., 2021). Thus, FL enables decentralized and collaborative training of a DL model in a privacy-preserving manner without the exchange of the actual data (Kairouz et al., 2021; Sheller et al., 2020). In addition, due to its emphasis on collaboration, FL opens up avenues for the creation of federated AI ecosystems to jointly solve DL tasks and improve the DL models and to ultimately create an added value for all participants in the ecosystem (Röder et al., 2021). Drawing on literature regarding the service-dominant logic and service innovation further substantiates the idea of leveraging federated AI ecosystems because value originates from the utilization and recombination of operant resources (e.g., knowledge) from multiple entities (Vargo and Lusch, 2004, 2008; Vargo et al., 2008). Furthermore, innovations do not emerge from isolation but rather from collaboration (Lusch and Nambisan, 2015). Therefore, participating in a federated AI ecosystem holds the potential for SME retail companies to sustain or regain their competitiveness and satisfy customers through new and innovative products and services.

Nevertheless, the design of an ecosystem—especially a federated AI ecosystem—can be very complex and challenging by itself as it involves a number of various structural, architectural and strategic decisions on multiple levels (Adner, 2017; Lusch and Nambisan, 2015). It is hence essential for both—researchers and practitioners—to be aware of the types of benefits expected from value co-creation within federated AI ecosystems and how these values relate to a participating company's size (Lusch and Nambisan, 2015). However, to the best of our knowledge, there are hardly any studies examining the possible business benefits of federated AI ecosystems, particularly in a retail context. Such research could assist practitioners in deciding whether it is worthwhile to establish or participate in a federated AI ecosystem and at the same time open up promising avenues for further research regarding the actual design of federated AI ecosystems.

To address this gap in the literature, our research investigates the value provided by federated AI ecosystems by the example of the retail industry. For this purpose, we compare three different types of models, namely a federated learning model, a centralized model, and multiple isolated models to solve a highly relevant AI task in the retail sector—the prediction of future sales for multiple retail stores (Ali, 2013; Huang et al., 2014; Na et al., 2019). Thereby, we aim to answer the following research question:

RQ *Which types of value are created for retail companies participating in a federated AI ecosystem for joint sales forecasting?*

Therefore, in the following sections, we first cover the conceptual foundations of our research and hereafter review related studies regarding FL in the area of the retailing industry and FL ecosystems to highlight the research gap. Subsequently, we describe our methodology along the principles put forward by Müller et al. (2016) for conducting a data science study to answer the formulated research question. Finally, we conclude our study and give an outlook on future research avenues.

2 Conceptual Foundations

2.1 Ecosystems

The ecosystem concept has its roots in biology, but was later transferred to the management and IS research discipline. An ecosystem can generally be described as a set of parties that interact while being mutually dependent on each other's capabilities in order to achieve a shared value proposition (Adner, 2017; Adner and Kapoor, 2010; Lusch and Nambisan, 2015; Pappas et al., 2018). Service ecosystems in particular originate from a service-dominant logic and provide the analytical basis for understanding how actors co-create value (Lusch and Nambisan, 2015; Trischler et al., 2020; Vargo and Lusch, 2017). From this perspective, the exchange of services rather than goods poses the foundation of any economic exchange with services referring to the utilization of own resources (e.g., knowledge) for the creation of value for oneself and others (Lusch and Nambisan, 2015; Vargo and Lusch, 2004, 2008; Vargo et al., 2015; Vargo and Lusch, 2017). Based on this rationale, a service ecosystem can be described as a largely self-contained and self-adjusting system consisting of loosely coupled actors that share a common institutional logic with the goal to generate added value for themselves and each other through the exchange of services (Lusch and Nambisan, 2015; Vargo and Lusch, 2014, 2016). To foster the exchange of services, Lusch and Nambisan (2015) suggest service platforms as a means to simplify the interaction between individual actors with their resources and thereby highlight the dual role of information technology as an operand and operant resource to enable and form value co-creation respectively.

2.2 Federated Learning

The concept of FL was first introduced by McMahan et al. (2017) and can be broadly described as a ML setting in which multiple parties (e.g., mobile devices or organizations) jointly solve a ML task without sharing their sensitive data (Kairouz et al., 2021). To achieve this goal, the parties join forces under the coordination of a neutral service provider (e.g., a university) to collaboratively train an ML model (Kairouz et al., 2021). Hereby, the data of the individual parties remains privately with the respective party and is not shared for training purposes (McMahan et al., 2017). To facilitate the learning task, the service provider initiates a ML model which is shared with all participating parties (McMahan et al., 2017). Subsequently, each party trains a local model based on private data to then send the model weights back to the service provider (Kairouz et al., 2021; McMahan et al., 2017). Once the service provider receives the local weights, they are aggregated to update the initially shared global model (McMahan et al., 2017). Finally, the service provider shares the updated global model with the participating parties and the process can be repeated all over again until an acceptable model performance is achieved (Kairouz et al., 2021; McMahan et al., 2017).

3 Related Work

Regarding the objective of the present study (i.e., to identify the types of value created within federated AI ecosystems by the example of the retail industry) related work can be broadly separated into two streams: (i) research on the application of FL in retail and (ii) studies concerning value creation within federated AI ecosystems. To collect the related literature, we conducted two literature reviews by loosely following the guidelines put forward by Vom Brocke et al. (2009). From the initial (i) 47[1] and (ii) 76[2] hits respectively, we sort out publications that do not meet the formal (e.g., books and patents) nor thematic requirements (i.e., (i) implementation in the retail sector; (ii) value creation within federated AI ecosystems) to ultimately obtain just seven relevant research articles.

In terms of FL applications in retail Ahmed et al. (2021b) develop a DL model utilizing FL for customer transaction classification for multiple domains (e.g., retail and healthcare). Similarly, Ahmed et al. (2021a, 2022) use the same DL model for collaborative customer clustering in commerce. Besides, Singh et al. (2021) employ another DL-based model (i.e., BERT) to determine customer compliance through social media data. However, when it comes to the evaluation, the aforementioned publications compare their

[1] Google Scholar search on November 8, 2021: *intitle:"federated learning" AND "retail“*

[2] Google Scholar search on November 8, 2021: *“federated Learning” AND ("ecosystem" OR "ecosystems") AND (“value creation” OR "value co creation" OR "value co-creation")*

respective models against other FL approaches, disregarding the added value compared to local approaches. Nevertheless, Singh et al. (2021) additionally report on the model performance in relation to a centralized model, revealing a slightly inferior performance of the FL approach.

Regarding value creation within federated AI ecosystems, we find this research stream to be in its infancy. This view is supported by the literature predominantly pointing towards the use of FL both inside and across enterprise boundaries to collaboratively build federated AI ecosystems (Fink et al., 2021; Röder et al., 2021). While the first-mentioned article is concerned with a rather conceptual view regarding the benefits of industrial AI ecosystems (Fink et al., 2021), the latter is directed towards federated AI ecosystems driven by edge intelligence in particular (Röder et al., 2021). In contrast, Idé and Raymond (2021) focus on blockchain and FL-based platforms for value and insights creation.

In sum, our review of the literature indicates that a significant research gap still exists regarding the benefits in terms of value generated through a federated AI ecosystem in general and in a retail context in particular. Consequently, the following sections focus on these benefits in more depth.

4 Research Approach

To address our research question, we conduct a data-driven sales forecasting in the context of the retailing industry. For the sake of transparency and reproducibility, we employ the methodology proposed by Müller et al. (2016) since these guidelines ensure traceability when carrying out data science projects within the field of information systems research. The steps comprise the phases of setting the research goals by anchoring the endeavor appropriately, (5.1) collecting the necessary data, (5.2) conducting the actual analysis, and (5.3) interpreting the results. Figure 1 illustrates this proceed.

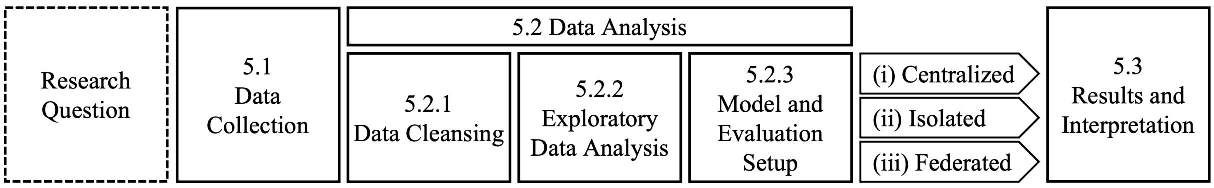

Figure 1. Research Approach in a Nutshell (based on Müller et al. (2016)).

To begin with, Müller et al. (2016) recommend adequately aligning the project with the central research question. As we are concerned with the analysis of the value of federated AI ecosystems on the retailing industry as opposed to rather traditional DL setups, we choose to compare three different types of setup scenarios—namely, (i) centralized, (ii) isolated, and (iii) federated. Through this, we firstly aim to elaborate on the value of privacy preservation by comparing (i) and (iii). Moreover, we highlight the value of collaboration within the scenario (iii), that is, the federated AI ecosystem opposed to the scenario (ii). Lastly, we investigate the associated generalization potential through collaboration (i.e., scenario (iii)) compared to no participation whatsoever (i.e., scenario (ii)). To this end, we make use of a publicly available dataset—again for the purpose of reproducibility—with historic sales data from multiple retail stores (cf. Section 5.1) to finally predict future sales for the upcoming month. However, prior to the actual analysis, we prepare the dataset through an extensive data cleansing phase (cf. Section 5.2.1) by rectifying the effects of previous extraordinary events (Fildes et al., 2006). Deeply interwoven with this step is the subsequent explanatory data analysis (5.2.2) which aims to better understand the data at hand (Tukey et al., 1977). Now, with the objective of comparing the various scenarios outlined above, we train multiple predictive DL-based models accordingly (cf. Section 5.2.3). Regarding this, we provide detailed information on the data partitioning for the various setups, the algorithms, and evaluation metrics were chosen to finally present the evaluation results obtained. Lastly, we interpret these results to shed light on the value added to the retail companies through collaboration within federated AI ecosystems (cf. Section 5.3).

5 Retail Case Example

5.1 Data Collection

In the retail sector, companies increasingly hold records about past sales. This data is frequently used to predict future sales and thereby fulfill demand timely without overfilling warehouses. However, this data is highly sensitive and critical to competition (Clifton and Marks, 1996). In addition, the DL-driven models to be investigated in the subsequent case example require large amounts of data for training. As a consequence, there is a scarcity of adequate data in the retail industry to conduct transparent and reproducible research. One publicly available dataset, however, provides sufficient data and is therefore well-suited for the overall research goal of this analysis. To be more precise, we use the openly accessible sales data supplied by the 1C Company data on Kaggle[3]. The dataset comprises sales data from 60 different retail shops covering 34 months from January 2013 to October 2015, totaling more than 2.9 million store-to-item sales pairs. In addition, further information about the items (e.g., name, category, and price), as well as the names of the stores are available.

5.2 Data Analysis

5.2.1 Data Cleansing and Preparation

In order to ensure data and model quality alike, an important step is to analyze the data for errors and inconsistencies (Rahm and Do, 2000; Zhang et al., 2003). Therefore, we conduct a preliminary descriptive analysis (cf. Table 1), which reveals that the dataset contains 7,356 negative item sales (e.g., returns) and 3,414 rather high daily item sales in single stores. Since we are only interested in predicting future sales, we omit these negative values. In addition, regarding the latter rather high values up to 2,169 item sales, it seems striking that the 99.9-th percentile equals 20. This is why we choose to omit all values above this percentile to finally obtain the cleaned up version of the dataset (Hodge and Austin, 2004). As mentioned before, our goal is to predict the monthly item sales at the store and item level. For this purpose, we sum up the monthly sales counts per item and shop for a more aggregated version of the data.

Statistic	**Minimum**	**Median**	**Mean**	**Maximum**	**Variance**	**Skewness**	**Kurtosis**
Value	*-22.0*	1.00	1.24	2169	6.85	272.83	177,477.79

Table 1. Descriptive Statistics.

5.2.2 Exploratory Data Analysis

To gain a deeper understanding for the aggregated data, we conduct an exploratory data analysis (Tukey et al., 1977). Therefore, we first explore the total item sales (cf. left-hand side of Figure 2). As a result, we found that monthly sales are subject to fluctuations and reach the annual peak around Christmas time. In addition to this annual trend, item sales continuously decrease over time which may be caused by the Russian economic crisis during this time span (Viktorov and Abramov, 2020). As these results indicate, overall item sales are influenced by both seasonality and trends. Thus, we decompose the time series to further investigate these effects for the item sales. The right-hand side of Figure 2 reveals that seasonality accounts for 2,060.31 of monthly item sales on average, while the underlying long-term sales trend is negative.

[3] Available at: https://www.kaggle.com/c/competitive-data-science-predict-future-sales

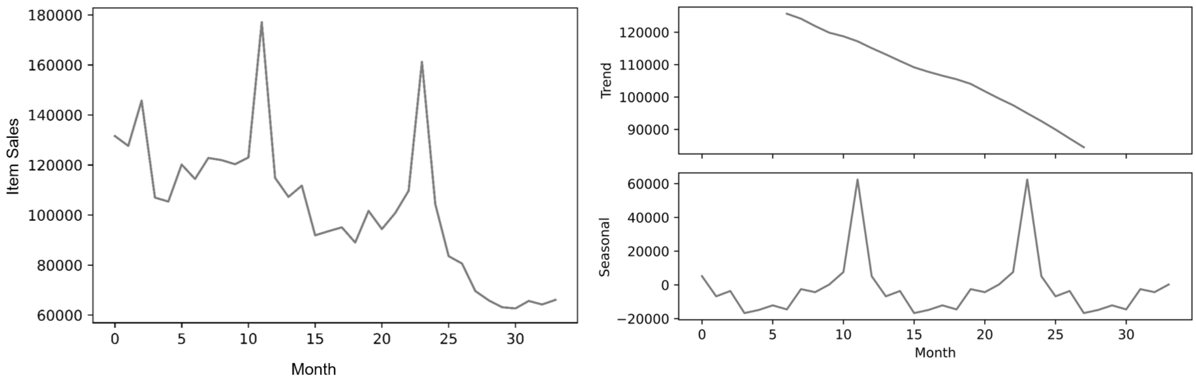

Figure 2. Overall and Decomposed Time Series for Item Sales.

By analyzing the sales on shop-level—with a few notable examples given in Figure 3—we find the total and monthly sales to differ extremely. Another remarkable aspect not depicted in Figure 3 is the rather high difference regarding the number of unique items per store.

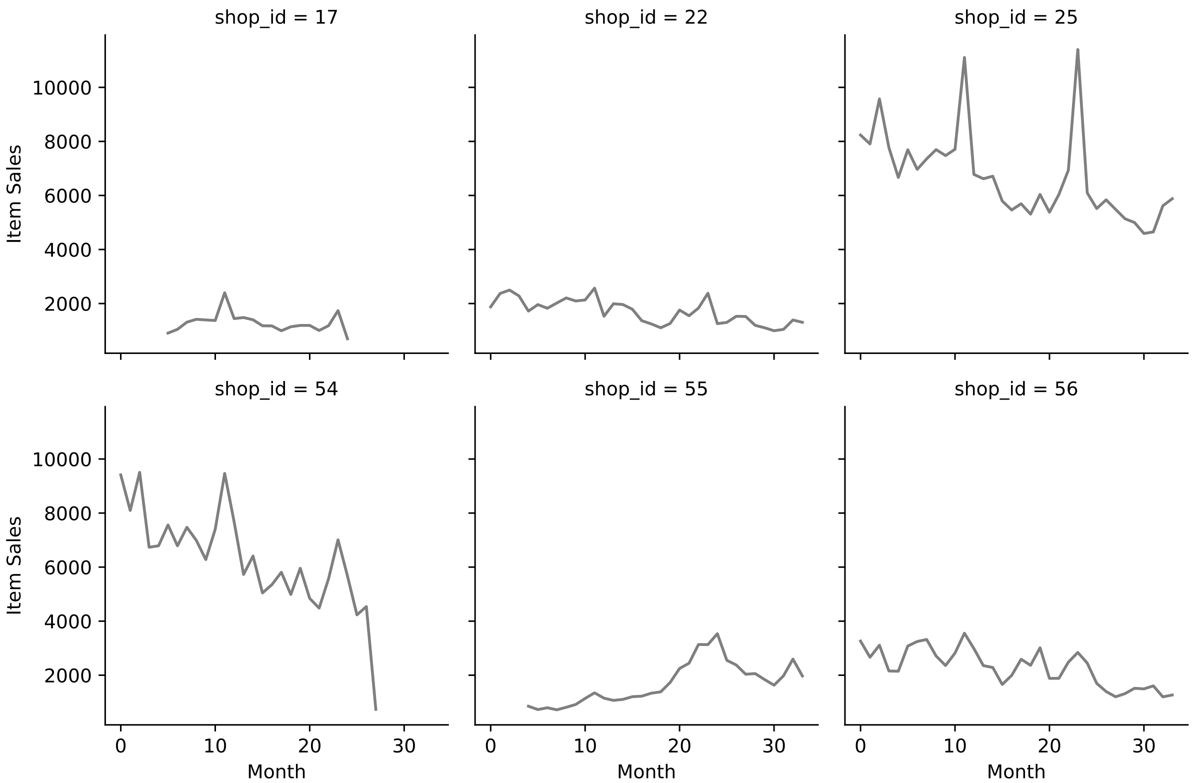

Figure 3. Monthly Sales per Shop.

Since not all shops did seamlessly record sales, we can broadly distinguish between three groups of stores. The first set of shops comprises of sales data over the total period of observation. We therefore refer to these shops as *consistently open*. The second group (i.e., *closed shops*) refers to stores with a lack of sales data at some point in the recording period until its end. Moreover, some shops only sold items during a specific time span. Thus, we assigned these stores to this group as well. Contrary, some shops just started recording item sales within the investigation period. We therefore declare these incidents as *recently opened stores*. Lastly, we notice some shops recording product sales for a certain period, followed by no sales for another one, only to have sales resumes afterwards. Based on this, we regard these stores as temporally closed (e.g., due to a major shop renovation), which is why we include them in the group of *recently opened stores*. The following Table 2 provides an overview of the number of shops per group—each with an average for the number of monthly items and unique items sold.

Shop Group	# Shops	Mean Monthly Sales	Mean Monthly Unique Sales
Constantly Opened Shops	32	2,275	1,042
Closed Shops	16	2,469	1,131
Recently Opened Shops	12	1,725	783

Table 2. Overview on Shop Groups.

A further examination of Figure 3 also demonstrates that across the entire observation period, some stores (e.g., store 25) consistently generated significantly higher sales than others (e.g., store 56). This suggests that—solely based on total sales—the stores may also differ significantly in their sizes. Since the total amount of sales is a popular and suitable proxy to infer the company size (Dang et al., 2018), we can divide the shops into three categories in terms of size. It is important to note that we only use the *constantly opened shops* for this purpose, as the lack of sales data for the other two shop groups (cf., Table 2) may lead to distortions (Dang et al., 2018). The first shop-size category consists of eight *small shops* with up to 43,291.50 total sales (i.e., 25.0-th percentile), whereas the eight *large shops* can each record a remarkable amount of more than 66,822.0 total sales (i.e., 75.0-th percentile). For the remaining 16 stores, the total number of sales remained within the 25- to 75-percentile range, so we grouped them as medium size stores. Table 3 provides further information regarding the sales distribution within the three shop-size groups. However, the minimum and maximum sales values indicated within a group do not necessarily meet the mentioned limits, as these were estimated based on the total sales of all *constantly opened shops*.

Shop Size Group	# Shops	Minimum Sales	Mean Sales	Maximum Sales
Small Shops	8	26,386.00	35,951.50	41,244.00
Medium Shops	16	43,974.00	54,443.87	66,414.00
Large Shops	8	68,046.00	121,745.25	234,274.00

Table 3. Overview Shop Size Groups of Constantly Opened Shops.

5.2.3 Model and Evaluation Setup

For the purpose of model evaluation, the choice of appropriate metrics is crucial (Gneiting, 2011). Regarding univariate time series forecasting, we can choose from a wide range of error-based measures. These include but are not limited to scale-dependent squared and absolute measures as well as scale-independent percentage measures—each with its respective advantages and disadvantages (Gneiting, 2011; Hyndman and Koehler, 2006). As the dataset under investigation contains rows with zero item sales, percentage-based metrics like the mean absolute percentage error are not suitable (Hyndman and Koehler, 2006). Contrastingly, the frequently employed scale-dependent error measures do not suffer from this disadvantage and are highly useful for the performance comparison of various methods (Hyndman and Koehler, 2006). Against this backdrop, we opt for two of these scale-dependent measures—namely, mean absolute error (MAE) and root mean squared error (RMSE). The metrics are accessed as follows:

$$MAE = (\frac{1}{n})\sum_{i=1}^{n}|y_i - x_i|$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i)^2}$$

Here, y_i represents the prediction for the i-th instance whereas x_i is the true value in this respect. The variable n stands for the total number of instances i to be compared.

As for the model deployed, time series forecasting tasks are increasingly facilitated by DL models. Among the rather popular models are convolutional neural networks (CNNs) and long short-term memory networks (LSTMs) since these models are capable to efficiently deal with noisy time series data (Fawaz et al., 2019; Livieris et al., 2020; Wang et al., 2017; Xue et al., 2019). Whereas CNNs are known for learning local trend features, LSTMs can capture long term dependencies (Xue et al., 2019). By combining both models, some authors like Livieris et al. (2020) demonstrated superior performance for time series predictions. Accordingly, we adopt the second CNN-LSTM type proposed by Livieris et al. (2020) for this study.

In order to explore the value created by federated AI ecosystems and thereby address the initially formulated research question, we train multiple CNN-LSTM models for the previously described scenarios—namely, (i) centralized, (ii) isolated, and (iii) federated—to evaluate their global, local and generalized predictive performance (Ek et al., 2021). Before training the models, we select the *constantly opened shops* (cf. Table 2) to then split this subset into a training (i.e., month 1 until 32), validation (i.e., month 2 until 33) and holdout set (i.e., month 3 until 34) to prevent information leakage (Hastie et al., 2009; Shmueli and Koppius, 2011). Thereby the last month of each set respectively represents the target value. In addition, to access the generalized predictive performance of the models, we employ the subset of the *recently opened shops* as a holdout-generalization set. Here, it should be noted, that we only use stores with item sales records that encompass more than just one month.

Since the training and evaluation setup for FL models can be thoroughly complex (Ek et al., 2021), we further elaborate on this in more detail in the following to ensure transparency. For all the models created, we use the Adam optimizer (Kingma and Ba, 2014)—even as the server optimizer for the FL model (Reddi et al., 2020).

Centralized Training. We train the centralized CNN-LSTM on the combined training set of all *constantly opened shops* over 50 epochs with a batch-size of 32. To access the predictive performance before the final evaluation, we monitor the model's performance for the combined validation set. Moreover, to prevent over-fitting, we employ the model's loss for the validation-set as an early stopping criterion. The centralized model converges after 24 epochs.

Isolated Training. To obtain the isolated models, we train a CNN-LSTM model for every shop from the *constantly opened shops* subset in separation given its respective training set—again over 50 epochs with a batch-size of 32. Likewise, we monitor the validation loss as described previously and utilize an early stopping rule to prevent over-fitting. Consequently, we receive 32 local models, which on average converge after 11 epochs.

Federated Training. As generally the case for FL, we set up a simulation environment consisting of one FL service provider and 32 locally dispersed clients (one per *constantly opened shop*)—with each one taking hold of its data to thus maintain privacy (Ek et al., 2021; Kairouz et al., 2021). The federated CNN-LSTM is subsequently trained collaboratively for 50 global rounds based on the aggregated model updates—weighted by the clients data quantity—from five local learning epochs per client. Once more, we use a batch size of 32 and monitor the loss for the validation set. Since FL models opposed to local models naturally require more time to converge (e.g., due to additional global epochs) in order to achieve good performance, we do not employ early stopping (Bonawitz et al., 2019). In line with this, the global model takes up to 35 epochs to converge. As we are interested in the effects of the above models regarding the global, local and generalizable predictive performance, we conduct three different types of evaluations. These are explained in the following.

Global Evaluation. Through this, we want to determine whether a FL model is capable to provide accurate global predictions. Furthermore, we want to access the extant as to which trade-offs exist between the (i) centralized and (iii) federated model. Hence, we evaluate these models performance for the test dataset of all *constantly opened shops*.

Local Evaluation. The local evaluation contributes to understanding the value created by FL for already participating organizations in a federated AI ecosystem and also highlights the trade-off towards centralized model training. To this end, we test the predictive power of the (i) centralized, (ii) isolated, and (iii) federated models against the individual test data of every *constantly opened shop* in separation. Here, each isolated model is benchmarked with its respective test data.

Generalized Evaluation. To assess the generalizability of the (i) centralized, (ii) isolated and (iii) federated models, we access their performance for the *recently opened shops*. Similarly to the local evaluation, each model is compared to the individual data per shop. Thereby, we highlight the value created through FL for new participants.

Evaluation Setting	Model	RMSE	MAE
Global Evaluation	Centralized	**0.61**	**0.17**
	Isolated	NA	
	Federated	0.62	0.20
Local Evaluation	Centralized	**0.56 ± 0.13**	**0.16 ± 0.04**
	Isolated	0.61 ± 0.14	0.21 ± 0.05
	Federated	0.57 ± 0.14	0.20 ± 0.04
Generalized Evaluation	Centralized	**1.19 ± 1.11**	**0.47 ± 0.51**
	Isolated	2.21 ± 1.09	1.46 ± 0.67
	Federated	*1.21 ± 1.13*	*0.50 ± 0.51*

Table 4. Evaluation Results.

According to the evaluation results (cf. Table 4), we find that the centralized model outperforms the other two for every evaluation setting. In terms of the global evaluation, the centralized model achieves an RMSE of 0.61 and is thereby slightly better compared to the federated model with 0.62. Note that the MAE values are shown for completeness purposes only as they do not represent the main evaluation criterion since we are searching for a reasonable prediction performance for outliers. Now, regarding the local evaluation setting, the centralized model again is barely better than the federated one with an mean RMSE of 0.56 opposed to 0.57. Likewise, the centralized model produces the least errors for the generalized evaluation setting. At first glance, it seems as if the models over-fit in the context of the generalized evaluation setting. Thus, we encourage to take a deeper look into the spread between the prediction errors. Consequently, we closely examine the descriptive statistics regarding the RMSE for each model in the following (cf. Table 5).

Evaluation Setting	Model	Minimum	25 % quantile	Median	75 % quantile	Maximum
Local Evaluation	Centralized	0.38	**0.46**	**0.55**	**0.59**	**0.96**
	Isolated	0.42	0.51	0.59	0.64	1.01
	Federated	**0.37**	0.47	0.57	**0.59**	0.99
Generalized Evaluation	Centralized	0.51	**0.55**	0.61	**1.16**	**3.98**
	Isolated	**0.48**	1.68	2.28	2.67	4.45
	Federated	0.52	0.56	**0.60**	1.19	4.06

Table 5. RMSE Distribution Generalized and Local Evaluation.

In this respect, we find the median RMSE values for the centralized (i.e., 0.55) and federated model (i.e., 0.57) to be approximately on the same level compared to the local evaluation. This allows us to conclude, that these models do not over-fit within the generalized evaluation setting. This also holds for the other statistics provided with the above table for these models. As for the isolated model, however, we notice a significant difference compared to the other two models—with the exception of the minimum value. Furthermore, we recognize a slightly lower median RMSE value for the federated model compared to the centralized model.

In summary, our results attribute superior performance to both—centralized and federated models—over the once trained in isolation. Additionally, we recognize a drastically worse forecasting capability for the local models opposed to the others with regards to *recently* or *reopened shops*. Counter to shops with small amounts of historical sales data, the centralized and federated models especially provide outstanding performance.

5.3 Result Interpretation and Discussion

According to Müller et al. (2016), the interpretation of the results represents the final step in a data science project. For this purpose, we now discuss our findings in light of existing theory and literature. We start with explaining the (i) *value of privacy preservation*. In the centralized setup, all data must be transferred and accumulated for global model training, whereas in the context of the FL setup the data are kept by the sovereign owners (Kairouz et al., 2021; McMahan et al., 2017; Yang et al., 2019). Bearing in mind the potential harm associated with open data sharing (Hann et al., 2007), FL offers a possible solution to data holders. However, as FL may affect the overall predictive performance, one has to carefully evaluate the worthiness of such an approach. In our study, we find the performance differences between the federated and centralized setups of being marginal (cf. Table 4). As these results are in line with previous work (e.g., Idé and Raymond (2021), McMahan et al. (2017), Sheller et al. (2020), and Yang et al. (2019)), we deduce that there is a performance trade-off between centralized and federated approaches based on privacy preservation. Consequently, the *value of privacy preservation* for parties within federated AI ecosystems can be determined by this trade-off through RMSE median comparison for example (cf. Figure 4). In this regard, according to the local evaluation, we can conclude that the *value of privacy preservation* accounts for a deviation of 2.09% of the median model performance.

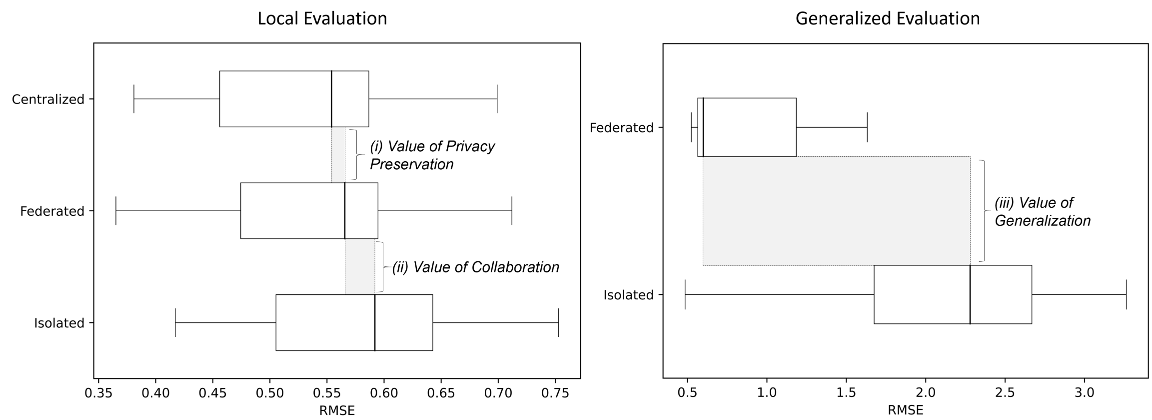

Figure 4. Values of Co-creation within Federated AI Ecosystems.

Next, we investigate the (ii) *value of collaboration*. The concept of FL is based on the idea, that multiple parties jointly solve a ML task, whereas in an isolated approach each party pursues this task on its own (Kairouz et al., 2021; McMahan et al., 2017; Sheller et al., 2020). Based on literature on value co-creation in the information systems research domain, a collaboration is successful if it is able to produce a specific outcome (Kotlarsky and Oshri, 2005). Since FL enables to collaboratively train a global model, we argue that the model is successful if it outperforms the local instances (Bogdanova et al., 2020). Regarding our analysis, we find the federated model to clearly overtake the isolated models on average (cf. Table 4), which again is consistent with previous related research (e.g., Sheller et al. (2020)). Thus, we infer that a federated AI ecosystem also holds a *value of collaboration*, which is also depicted in Figure 4 and results in a median model performance difference of 4.32% for our example.

Lastly, we shed light on the (iii) *value of generalizability*. In general, DL models require a vast amount of data in order to achieve generalizability (Sandfort et al., 2019). In this regard, the principle of FL is based on implicit data utilization, whereas the isolated models can only rely on their respective dataset. Following the literature on the service-dominant logic through an ecosystems perspective (Lusch and Nambisan, 2015; Vargo and Lusch, 2004, 2008; Vargo et al., 2015; Vargo and Lusch, 2017), which ascribes a dual role to each actor in an ecosystem (i.e., consuming and producing), we argue that a FL model can benefit from the shared knowledge (i.e., in the form of model weights) through the inclusion of new clients while these clients in turn benefit equally from the FL model. In this context our study reveals that the federated setting provides a better generalization performance for completely new retail shops in comparison to the isolated models (cf. Table 4). This behavior corresponds to the results of Ek et al. (2021). We therefore conclude that a federated AI ecosystem further holds the *value of generalizability* for both—the ecosystem itself and new participants. Again, this value can be measured by comparing the performance gap between the federated and isolated approaches based on the median RMSE prediction errors, which turns out to be quite impressive by a difference of 280% (cf. Figure 4).
In this context, it is of particular interest to examine how these values—namely, the *value of privacy preservation*, the *value of collaboration* and the *value of generalizability*—relate to the size of a company (cf. Table 3). In our example, the federated averaging algorithm used to aggregate the client updates employs a weighting based on the data quantity used for training (McMahan et al., 2017). Thus, we assume that the proposed values within federated AI ecosystems are equally tied to the store size, as the model is more adjusted to the stores that provide more data for the training procedure. With this in mind, we now examine the results of the local evaluation at the shop size level (cf. Figure 5). We find that the *value of privacy preservation* decreases slightly with increasing company size (i.e., median performance difference: small shops 2.83%, medium-sized shops 2.39% and large shops 2.12%). This indicates that smaller companies may face a greater privacy-performance trade-off by participating in a federated AI ecosystem opposed to larger companies. With regard to the comparison between isolated and federated model training, the *value of collaboration* does not increase linearly with the shop size. Interestingly, although the median performance delta between medium-sized shops (i.e., 6.52%) and large shops (i.e., 6.94%) is rather small, a considerable drop can be observed for the small shops (i.e., 4.25%). Accordingly, it is evident that larger as well as medium-sized companies obtain a greater *value of collaboration*. Hence, this value is comparatively limited for small companies in the federated AI ecosystem. Therefore, we conclude that the values a company can derive directly from participating in a federated AI ecosystem (i.e., privacy preservation and collaboration) depend on the resources (i.e., data) the company itself is willing or able to contribute to the ecosystem.

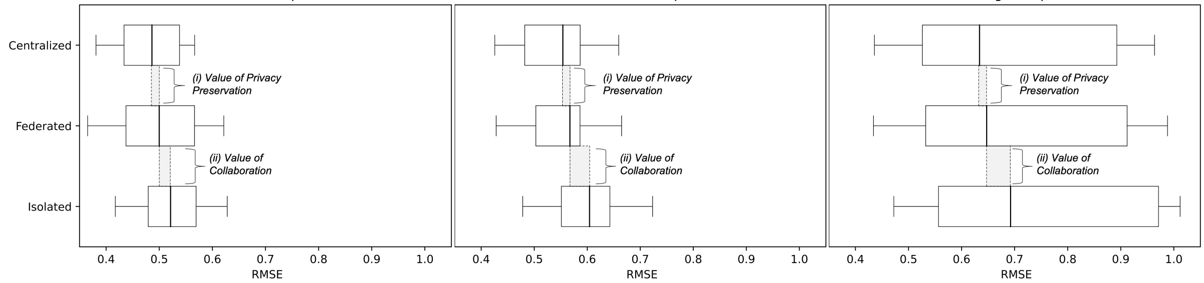

Figure 5. Value of Privacy Preservation and Collaboration per Shop Size.

Regarding the *value of generalization*, Figure 6, in contrast, depicts a different picture. Accordingly, the participation of small shops in a federated AI ecosystem yields to an astonishing generalization performance gain of 627.20%, while the performance gains for large shops are similarly remarkable with 302.90%. Interestingly, the performance increases for medium-sized companies are comparatively low at

64%. Thus, we draw the conclusion, that the benefit for a company from participating in a federated AI ecosystem may not solely be determined by the resources (i.e., data) it can contribute to the ecosystem. Since medium-sized companies represent half of the stores in our sample, whereas small and large stores each represent a fourth (cf., Table 3), we presume that the *value of generalization* for a company may be determined by its disparity from the prevailing client characteristics and data distribution within a federated AI ecosystem.

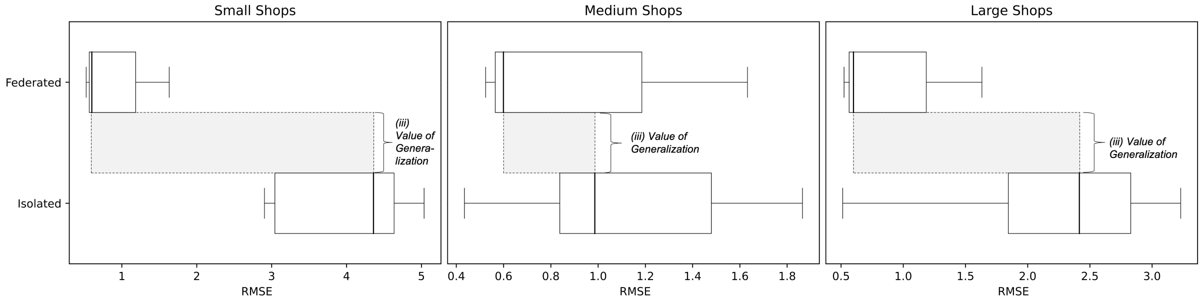

Figure 6. Value of Generalizability per Shop Size.

6 Conclusion and Outlook

Advanced DL technologies offer great potential for organizations and evermore outperform the rather traditional AI methods (e.g., ML). However, to fully unleash this value, DL algorithms require vast amounts of data in the first place to achieve a sufficiently high performance. Since companies—especially SMEs or novel market entrants—do not possess such rich data resources, the promised value associated with DL often remains untapped. This disadvantage may have far-reaching consequences and could ultimately diminish or even take away the market position of these companies. One way to counteract this adverse scenario for companies is to collaborate by data sharing. Yet, due to privacy concerns, this approach is rarely chosen. Therefore, this article highlights the emerging concept of FL to provide organizations with the option to collaboratively train DL models through weight sharing (i.e., the federated AI ecosystem) without the sacrifice of privacy violations. Because the setup and coordination of such a system can be thoroughly complicated, it is crucial for responsible decision-makers to communicate the expected benefits accordingly—which are as yet largely unclear. Therefore, we investigate these effects for the case of a multi-store retailer predicting item sales. As a result, we identified three kinds of values, namely *collaboration*, *privacy preservation*, and finally *generalization* and demonstrated how these values are tied to the size of a company.

Regarding the limitations, first and foremost the present analysis is dependent on the specific implementation, that is the use case (i.e., retail dataset), preprocessing steps, algorithms, evaluation metrics, DL model architecture, and FL averaging method. Emphasizing the latter point, a different weighting policy (e.g., uniformly over all clients—opposed to data volume dependent) could lead to other results (Kairouz et al., 2021). In particular, given the relationship between the *value of collaboration* and company size, it seems reasonable to assume that a uniform weighting policy may boost small companies participating in the federated AI ecosystem, whereas the opposite could be true for larger companies. In contrast, the *value of privacy-preservation* reveals minor variations due to company size, which may also apply for different weighting policies. Likewise, the *value of generalizability* may also be influenced by another weighting policy and another server optimizer might change the forecast as well (Reddi et al., 2020). Furthermore, the practical utility of federated AI ecosystems and its inherent value are limited by the training and convergence time of the FL model. Since in our example the FL model requires the most epochs to converge, there is a clear trade-off between convergence time and performance. This is particularly noticeable for the isolated models of small companies, as they benefit proportionally the least from participating in a federated AI ecosystem—apart from generalizability— but require disproportionately more resources (e.g., computational resources per training epoch) compared to isolated training. Therefore, despite the generated values, participation in a federated AI ecosystem might not always be economically viable for specific companies.

Despite the aforementioned limitations, our study holds valuable implications for both—researchers and practitioners. With regard to the former, we connect the research streams of FL with the service-dominant logic and ecosystems to explore the values created or co-created though the participation in a federated AI ecosystem. Thereby, we contribute to the body of knowledge through three explored types of values (e.g., the values of *collaboration*, *privacy preservation*, and *generalization*) and their association to the companies size. In particular, we highlight that, both—the *value of collaboration* and the *value of privacy preservation*—are positively influenced to the companies size, while the *value of collaboration* a company can capture may be driven by it's variation from predominant client sizes and data distributions within a federated AI ecosystem. These values can easily be transferred to other domains and applications—aside from the retail sector—to encourage participation within the context of a federated AI ecosystem. With regards to the practical implications, we first attribute a high utility to FL for the task of sales forecasting in retail. Here, the potentially complex process of administration (e.g., architectural decisions) can be shifted to the central service provider and thereby lower the barriers to entry for new collaborators to the federated AI ecosystem. Moreover, our analysis reveals, that companies of different size mutually benefit from each other through the collaboration within an federated AI ecosystem, which makes it apparent for practitioners that the joint formation in such ecosystems is a valuable and effective approach, regardless the size of the company. In addition, we attribute high performance to CNN-LSTM models for time series forecasting.

In conclusion, our paper opens up promising avenues for further research concerned with federated AI ecosystems. To this end, we first want to encourage to investigate and verify the effects for other datasets, domains (e.g., finance or healthcare) and likewise other tasks (e.g., fraud detection or predictive maintenance). Second, since the performance of the FL model seems to decrease slightly on our generalization data, the question arises how the inclusion of the respective additional clients may affect the overall performance of a federated AI ecosystem. As this type of question is a prime example of the network effects research stream (Economides, 1996; Katz and Shapiro, 1994; Kumar et al., 2021; Parker and Van Alstyne, 2005), a fruitful future work could investigate how the total number of participating clients or their respective data quality affect the values of a federated AI ecosystem. Thereby, such a research could contribute to a deeper understanding regarding the interactive nature of value co-creation within a federated AI ecosystem. Third, it is relevant to clarify what motivates or hinders companies to participate in a federated AI ecosystem. To this end, future work grounded in qualitative or quantitative research approaches could examine the factors that influence entrepreneurial decisions to adopt and participate in such ecosystems. Furthermore, the questions of how to design an ecosystem and a platform respectively to enable and enhance value co-creation through FL and the role of the service provider for the federated AI ecosystem could be analyzed more in depth as well. In addition, it would be useful to investigate how to design more appropriate metrics for performance comparison between centralized, federated, and local models, since for example training times, resource utilization, and network-related factors may be considered as well (Kairouz et al., 2021). Moreover, to overcome reservations on the client side due to a lack of traceability regarding the FL models predictions, trust-enabling methods should be investigated thoroughly—specifically for federated AI ecosystems as well.

References

Adadi, A. and M. Berrada (2018). "Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI)." *IEEE Access* 6, 52138–52160.

Adner, R. (2017). "Ecosystem as structure: An actionable construct for strategy." *Journal of management* 43 (1), 39–58.

Adner, R. and R. Kapoor (2010). "Value creation in innovation ecosystems: how the structure of technological interdependence affects firm performance in new technology generations." 31 (3), 306–333.

Ahmed, U., J. C.-W. Lin, and G. Srivastava (2021a). "Deep-Attention Model to Analyze Reliable Customers via Federated Learning," 1–7.

Ahmed, U., G. Srivastava, and J. C.-W. Lin (2021b). "A Federated Learning Approach to Frequent Itemset Mining in Cyber-Physical Systems." 29 (4).

Ahmed, U., G. Srivastava, and J. C.-W. Lin (2022). "Reliable customer analysis using federated learning and exploring deep-attention edge intelligence." *Future Generation Computer Systems* 127, 70–79.

Ali, G. (2013). "Driver moderator method for retail sales prediction." *International Journal of Information Technology & Decision Making* 12 (06), 1261–1286.

Bauer, M., C. van Dinther, and D. Kiefer (2020). "Machine Learning in SME: An Empirical Study on Enablers and Success Factors," 1–10.

Bengio, Y., I. Goodfellow, and A. Courville (2017). *Deep learning*. Vol. 1. MIT press Massachusetts, USA.

Bogdanova, A., A. Nakai, Y. Okada, A. Imakura, and T. Sakurai (2020). "Federated learning system without model sharing through integration of dimensional reduced data representations." eprint: arXiv:2011.06803.

Bonawitz, K. A., H. Eichner, W. Grieskamp, D. Huba, A. Ingerman, V. Ivanov, C. M. Kiddon, J. Konečný, S. Mazzocchi, B. McMahan, T. V. Overveldt, D. Petrou, D. Ramage, and J. Roselander (2019). "Towards Federated Learning at Scale: System Design."

Chui, M., J. Manyika, M. Miremadi, N. Henke, R. Chung, P. Nel, and S. Malhotra (2018). "Notes from the AI frontier: Applications and value of deep learning." *McKinsey Global Institute*. URL: https://www.mckinsey.com/~/media/mckinsey/featured%20insights/artificial%20intelligence/notes%20from%20the%20ai%20frontier%20applications%20and%20value%20of%20deep%20learning/notes-from-the-ai-frontier-insights-from-hundreds-of-use-cases-discussion-paper.pdf.

Clifton, C. and D. Marks (1996). "Security and privacy implications of data mining," 15–19.

Dang, C., Z. F. Li, and C. Yang (2018). "Measuring firm size in empirical corporate finance." *Journal of banking & finance* 86, 159–176.

Economides, N. (1996). "The economics of networks." *International journal of industrial organization* 14 (6), 673–699.

Ek, S., F. Portet, P. Lalanda, and G. Vega (2021). "A federated learning aggregation algorithm for pervasive computing: Evaluation and comparison."

Fawaz, H. I., G. Forestier, J. Weber, L. Idoumghar, and P.-A. Muller (2019). "Deep learning for time series classification: a review." *Data mining and knowledge discovery* 33 (4), 917–963.

Fildes, R., P. Goodwin, and M. Lawrence (2006). "The design features of forecasting support systems and their effectiveness." *Decision Support Systems* 42 (1), 351–361.

Fink, O., T. Netland, and S. Feuerriegel (2021). *Artificial intelligence across company borders*. eprint: arXiv:2107.03912.

Gneiting, T. (2011). "Making and Evaluating Point Forecasts." *Journal of the American Statistical Association* 106 (494), 746–762.

Grewal, D., D. K. Gauri, A. L. Roggeveen, and R. Sethuraman (2021). "Strategizing Retailing in the New Technology Era." *Journal of Retailing* 97 (1). Re-Strategizing Retailing in a Technology Based Era, 6–12.

Grewal, D., A. L. Roggeveen, and J. Nordfält (2017). "The future of retailing." *Journal of retailing* 93 (1), 1–6.

Guha, A., D. Grewal, P. K. Kopalle, M. Haenlein, M. J. Schneider, H. Jung, R. Moustafa, D. R. Hegde, and G. Hawkins (2021). "How artificial intelligence will affect the future of retailing." *Journal of Retailing* 97 (1), 28–41.

Hann, I.-H., K.-L. Hui, S.-Y. T. Lee, and I. P. Png (2007). "Overcoming Online Information Privacy Concerns: An Information-Processing Theory Approach." *Journal of Management Information Systems* 24 (2), 13–42.

Hastie, T., R. Tibshirani, and J. Friedman (2009). *The Elements of Statistical Learning. Data Mining, Inference, and Prediction*. 2nd Edition. Springer Series in Statistics. New York: Springer.

Hodge, V. and J. Austin (2004). "A Survey of Outlier Detection Methodologies." *Artificial Intelligence Review* 22 (2), 85–126.

Huang, T., R. Fildes, and D. Soopramanien (2014). "The value of competitive information in forecasting FMCG retail product sales and the variable selection problem." *European Journal of Operational Research* 237 (2), 738–748.

Hyndman, R. J. and A. B. Koehler (2006). "Another look at measures of forecast accuracy." *International journal of forecasting* 22 (4), 679–688.

Idé, T. and R. Raymond (2021). "Decentralized Collaborative Learning with Probabilistic Data Protection," 234–243.

Jöhnk, J., M. Weißert, and K. Wyrtki (2021). "Ready or Not, AI Comes—An Interview Study of Organizational AI Readiness Factors." *Business & Information Systems Engineering* 63 (1), 5–20.

Kairouz, P., H. B. McMahan, B. Avent, A. Bellet, M. Bennis, A. N. Bhagoji, K. Bonawitz, Z. Charles, G. Cormode, R. Cummings, R. G. L. D'Oliveira, H. Eichner, S. E. Rouayheb, D. Evans, J. Gardner, Z. Garrett, A. Gascón, B. Ghazi, P. B. Gibbons, M. Gruteser, Z. Harchaoui, C. He, L. He, Z. Huo, B. Hutchinson, J. Hsu, M. Jaggi, T. Javidi, G. Joshi, M. Khodak, J. Konecný, A. Korolova, F. Koushanfar, S. Koyejo, T. Lepoint, Y. Liu, P. Mittal, M. Mohri, R. Nock, A. Özgür, R. Pagh, H. Qi, D. Ramage, R. Raskar, M. Raykova, D. Song, W. Song, S. U. Stich, Z. Sun, A. T. Suresh, F. Tramèr, P. Vepakomma, J. Wang, L. Xiong, Z. Xu, Q. Yang, F. X. Yu, H. Yu, and S. Zhao (2021). "Advances and Open Problems in Federated Learning." *Foundations and Trends in Machine Learning* 14 (1–2), 1–210.

Katz, M. L. and C. Shapiro (1994). "Systems competition and network effects." *Journal of economic perspectives* 8 (2), 93–115.

Kingma, D. P. and J. Ba (2014). *Adam: A Method for Stochastic Optimization*. eprint: arXiv:1412.6980.

Kotlarsky, J. and I. Oshri (2005). "Social ties, knowledge sharing and successful collaboration in globally distributed system development projects." *European Journal of Information Systems* 14 (1), 37–48.

Kumar, A., B. Finley, T. Braud, S. Tarkoma, and P. Hui (2021). "Sketching an ai marketplace: Tech, economic, and regulatory aspects." *IEEE Access* 9, 13761–13774.

Livieris, I. E., E. Pintelas, and P. Pintelas (2020). "A CNN–LSTM model for gold price time-series forecasting." 32 (23), 17351–17360.

Lusch, R. F. and S. Nambisan (2015). "Service innovation." *MIS Quarterly* 39 (1), 155–176.

Mahmoud, A. B., S. Tehseen, and L. Fuxman (2020). "The Dark Side of Artificial Intelligence in Retail Innovation." *Retail Futures*, 165–180.

McMahan, B., E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas (2017). "Communication-efficient learning of deep networks from decentralized data," 1273–1282.

Müller, O., I. Junglas, J. Vom Brocke, and S. Debortoli (2016). "Utilizing big data analytics for information systems research: challenges, promises and guidelines." *European Journal of Information Systems* 25 (4), 289–302.

Na, L., C. Ying, and B. De (2019). "The Empirical Analysis of Convenience Store Sales Forecast Based on Multidimensional Data in Emerging Markets." ICBIM '19, 32–36.

Najafabadi, M. M., F. Villanustre, T. M. Khoshgoftaar, N. Seliya, R. Wald, and E. A. Muharemagic (2014). "Deep learning applications and challenges in big data analytics." *Journal of Big Data* 2, 1–21.

Oosthuizen, K., E. Botha, J. Robertson, and M. Montecchi (2021). "Artificial intelligence in retail: The AI-enabled value chain." *Australasian Marketing Journal* 29 (3), 264–273.

Pappas, I. O., P. Mikalef, M. N. Giannakos, J. Krogstie, and G. Lekakos (2018). "Big data and business analytics ecosystems: paving the way towards digital transformation and sustainable societies." 16 (3), 479–491.

Parker, G. G. and M. W. Van Alstyne (2005). "Two-sided network effects: A theory of information product design." *Management science* 51 (10), 1494–1504.

Rahm, E. and H. H. Do (2000). "Data cleaning: Problems and current approaches." *IEEE Data Eng. Bull.* 23 (4), 3–13.

Reddi, S., Z. Charles, M. Zaheer, Z. Garrett, K. Rush, J. Konečný, S. Kumar, and H. B. McMahan (2020). *Adaptive Federated Optimization*. eprint: arXiv:2003.00295.

Röder, M., P. Kowalczyk, and F. Thiesse (2021). "Leveraging the Potentials of Federated AI Ecosystems." Ed. by F. Ahlemann, R. Schütte, and S. Stieglitz, 61–68.

Samek, W. and K. Müller (2019). "Towards Explainable Artificial Intelligence." In: *Explainable AI: Interpreting, Explaining and Visualizing Deep Learning*. Ed. by W. Samek, G. Montavon, A. Vedaldi, L. K. Hansen, and K.-R. Müller. Cham: Springer International Publishing, pp. 5–22.

Sandfort, V., K. Yan, P. J. Pickhardt, and R. M. Summers (2019). "Data augmentation using generative adversarial networks (CycleGAN) to improve generalizability in CT segmentation tasks." *Scientific Reports* 9 (1).

Schmidt, J., M. R. Marques, S. Botti, and M. A. Marques (2019). "Recent advances and applications of machine learning in solid-state materials science." *npj Computational Materials* 5 (1), 1–36.

Shankar, V., K. Kalyanam, P. Setia, A. Golmohammadi, S. Tirunillai, T. Douglass, J. Hennessey, J. Bull, and R. Waddoups (2021). "How Technology is Changing Retail." *Journal of Retailing* 97 (1), 13–27.

Sheller, M. J., B. Edwards, G. A. Reina, J. Martin, S. Pati, A. Kotrotsou, M. Milchenko, W. Xu, D. Marcus, R. R. Colen, et al. (2020). "Federated learning in medicine: facilitating multi-institutional collaborations without sharing patient data." *Scientific reports* 10 (1), 1–12.

Shmueli and Koppius (2011). "Predictive Analytics in Information Systems Research." 35 (3), 553.

Singh, A., T. Sen, S. Saha, and M. Hasanuzzaman (2021). "Federated Multi-Task Learning for Complaint Identification from Social Media Data." HT '21, 201–210.

Trischler, J., M. Johnson, and P. Kristensson (2020). "A service ecosystem perspective on the diffusion of sustainability-oriented user innovations." *Journal of Business Research* 116, 552–560.

Tukey, J. W. et al. (1977). *Exploratory data analysis*. Vol. 2. Reading, Mass.

Ulrich, P. S., V. Frank, and M. Kratt (2021). "Adoption of Articial Intelligence in German SMEs–Results from an Empirical Study," 1.

Vargo, S. L. and R. F. Lusch (2004). "Evolving to a new dominant logic for marketing." *Journal of marketing* 68 (1), 1–17.

Vargo, S. L. and R. F. Lusch (2008). "Service-dominant logic: continuing the evolution." *Journal of the Academy of marketing Science* 36 (1), 1–10.

Vargo, S. L. and R. F. Lusch (2014). "Inversions of service-dominant logic." *Marketing theory* 14 (3), 239–248.

Vargo, S. L. and R. F. Lusch (2016). "Institutions and axioms: an extension and update of service-dominant logic." *Journal of the Academy of marketing Science* 44 (1), 5–23.

Vargo, S. L., P. P. Maglio, and M. A. Akaka (2008). "On value and value co-creation: A service systems and service logic perspective." *European management journal* 26 (3), 145–152.

Vargo, S. L., H. Wieland, and M. A. Akaka (2015). "Innovation through institutionalization: A service ecosystems perspective." *Industrial Marketing Management* 44, 63–72.

Vargo, S. L. and R. F. Lusch (2017). "Service-dominant logic 2025." *International Journal of Research in Marketing* 34 (1), 46–67.

Viktorov, I. and A. Abramov (2020). "The 2014–15 Financial Crisis in Russia and the Foundations of Weak Monetary Power Autonomy in the International Political Economy." *New Political Economy* 25 (4), 487–510.

Vom Brocke, J., A. Simons, B. Niehaves, B. Niehaves, K. Reimer, R. Plattfaut, and A. Cleven (2009). "Reconstructing the Giant: On the Importance of Rigour in Documenting the Literature Search Process." *17th European Conference on Information Systems (ECIS 2009).*

Wang, Z., W. Yan, and T. Oates (2017). "Time series classification from scratch with deep neural networks: A strong baseline," 1578–1585.

Xue, N., I. Triguero, G. P. Figueredo, and D. Landa-Silva (2019). "Evolving deep CNN-LSTMs for inventory time series prediction," 1517–1524.

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

Zhang, S., C. Zhang, and Q. Yang (2003). "Data preparation for data mining." *Applied Artificial Intelligence* 17 (5-6), 375–381.