AI for Impairment Accounting

Sören Hartung and Manuela Führer

1 Introduction

1.1 Initial Situation

In the euro zone, the interest rates are still at a historically low level. Regulation, driven by the last financial crisis, puts additional cost pressure on the banks. The traditional banks must face agile competitors (Fintech companies) and technologically strong and well capitalized companies (BigTech[1]) in the retail banking segment.

With serious threats to the revenue potential of certain segments, the traditional banks seek to improve their offerings and evaluate where to save costs by using automation (like RPA) and artificial intelligence to optimize processes.

[1] GAFAM (Google, Amazon, Facebook, Apple and Microsoft) and BATX (Baidu, Alibaba, Tencent and Xiaomi).

S. Hartung · M. Führer (✉)
Helaba, Frankfurt, Germany
e-mail: Manuela.Fuehrer@helaba.de

V. Liermann and C. Stegmann (eds.), *The Digital Journey of Banking and Insurance, Volume I*, https://doi.org/10.1007/978-3-030-78814-8_6

The imperative of efficient improvement is there; the banks recognize that cost cutting needs to be applied in an intelligent manner: by optimizing processes and keeping the functionality required.[2]

1.2 Bank-Wide Setup

For Helaba, data analytics is one facet of digitalization. In Fig. 1, this facet is broken down into six sub-areas. These areas are sorted by topics. There are sections with a strong client focus, such as text mining or segmentation, and units with a more internal concentration, such as anomaly detection and forecasting. Classification and frequent patterns are applied in both (client-related and internal) areas.

Helaba identifies use cases for digitalization with a well-orchestrated process. The use cases are always driven by the business units, solving a problem formulated by the business department.

Figure 2 shows the three main steps in use case selection and the time-frame given for the different steps. The data innovation workshop took place in March 2019 and 13 departments participated in the workshop. After selecting three use cases, the feasibility was tested within the framework of POCs and evaluated. The use cases with a positive evaluation were then implemented in a productive environment.

The basic concept for the underlying process has been to generate and discuss multiple ideas with a diverse and cross-functional group of internal experts, test these ideas in the shortest possible time and with as few resources as possible and to bring only the ideas that are worth the effort into the operationalization phase.

The second phase took around eight months.[3] The two success factors in this phase were getting open-minded colleagues on board who are eager to work on innovative solutions for day-to-day problems, and getting the commitment of senior management to try new solutions even if there is a high probability that the immediate results are not good. One of the POCs that actually delivered very good results is described in detail in the following sections.

[2] The dysfunctional cost cutting would be to brute force reduction of needed functionality.

[3] The Proof of Concept (POC) only took eight weeks.

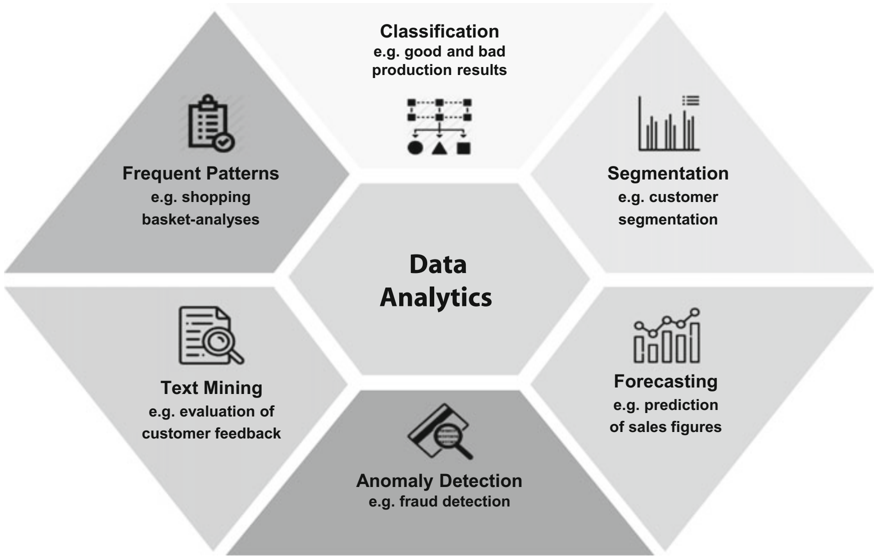

Fig. 1 Data analytics (© Helaba)

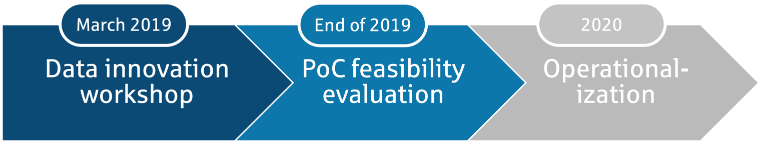

Fig. 2 Overall identification process (© Helaba)

1.3 Situation—Accounting Department

Generating financial statements in the accounting department means collecting huge amounts of data from several sources, and unavoidably some of the data is incorrect. The textbook approach would be: identify the errors, adjust the data delivery, customize the calculation software itself. The reality looks different. Identification of errors is not that simple and means manual workload. Certain errors only occur in one month and are gone the next. Some errors can only be solved through very high investments in technical and organizational process flows. However, this would not always be efficient from a cost/benefit perspective. Sometimes, adjusting the error after the process is more efficient than adjusting the process that produces good results for 99.9% of the cases.

1.4 Structure of the Chapter

In Sect. 2, the business requirements are discussed. The chosen technical setup of the POC is described in Sect. 3. The approach taken in the project (POC) is delivered in Sect. 4 and the results and the lessons learned are shown in Sect. 5. Section 6 sums the chapter up and provides an outlook of future developments.

2 Business Setting

2.1 Challenges

The bank has implemented decentralized responsibility for accounting material and data quality. This means that the recording of data to ensure correct posting processes and business transactions as well as the monitoring of accounts is carried out locally in the product areas, because knowledge of the economic facts is concentrated there. Nevertheless, the accounting department has a control function through its overall responsibility for the financial statements and the internal control system. For impairment accounting, approximately 80,000 transactions per month are performed and processed for posting risk provisions using >15 systems. The process for preparing the financial statements is subject to tight time restrictions.

Every month, approximately 5000 anomalies occur in the processing chain, about 50 of which must be corrected manually every month due to their criticality. The various causes of errors include data inconsistency, time lags in the provision of data, technical restrictions in the system landscape and the generally sensitive and complex processing chain (Fig. 3).

2.2 Correction Process

Today's error correction process is structured as follows: two employees in the accounting department identify and analyze error constellations every month after the last day of business. In this largely manual process, three Microsoft Excel-based applications are currently still in use. Errors are either forwarded to product areas or to IT operations for correction or corrected by the accounting department independently. Since errors often only become apparent late in the processing, they can sometimes only be corrected in

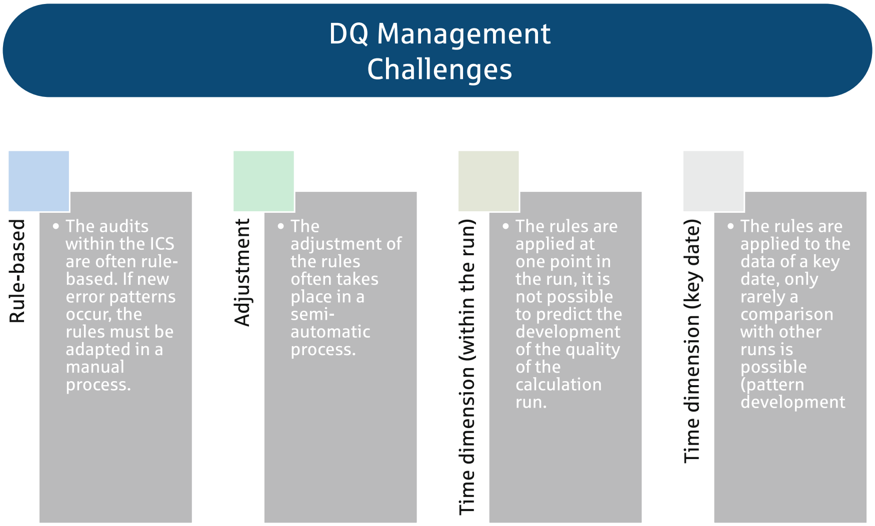

Fig. 3 Challenges in data quality management (© Helaba)

subsequent systems. Other departments do not benefit from these correction entries and rely on an incorrect database. The total expenditure for the outlined process is about 20 person days/month.

2.3 POC Target

In general, POCs aim to validate a technology or show the feasibility of an approach. This specific POC aims—in addition—to generate practical experience with data analysis and machine learning. This experience should enable the department to validate other subjects in terms of their automation potential.

The main goals of the POC are shown in Fig. 4. The major objective is to identify error patterns and improve the data quality. The identification of anomalies by using machine learning can improve the time to correct the errors. This verification with machine learning models can be established as an alternative to the existing rule-based approach.

A second objective is to review the appropriate machine learning models and their performance in terms of the challenge.

Furthermore, the POC also targets the knowledge build-up and the important practical experience with the models and tools.

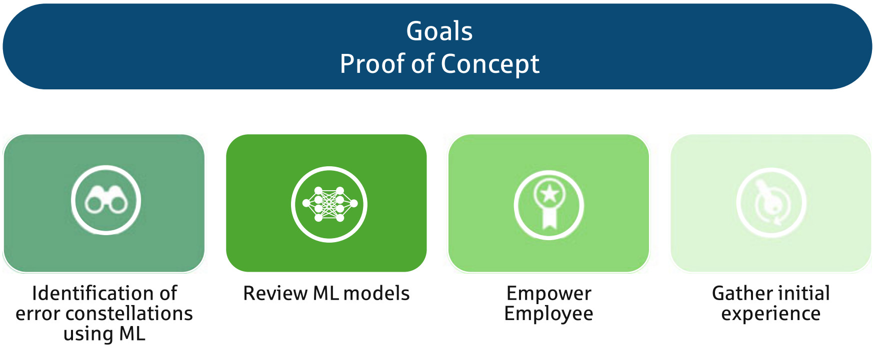

Fig. 4 Goals of the POC (© Helaba)

3 Technical Setup—Proof of Concept

3.1 Data Flow

The POC analyzes the data flows in the context of processing the impairments in the accounting department. For this purpose, the data from the upstream systems and the systems used to determine the amount of the impairment are compared with the values in reporting or consolidation. If there are deviations, an error has occurred (ICS rule).[4]

In Fig. 5, the landscape of the systems involved is presented. The data flow is from bottom to top. The data is extracted from the impairment system[5] and the operative systems (core systems). In a data warehouse, the data is harmonized and can then be transferred to the accounting subledger systems (SAP AFI). Afterward, the data is passed on to the reporting system and further to the consolidation system.

One rule (among many other rules) checks if the sum of the key figures used for the impairment value compilation (in the reporting system) is still equal to the original values in the impairment system. In terms of machine learning, we call the result of the comparison a label (the values the model should predict). In the POC flow, we decided to differentiate based on the severity of the deviation (impact sensitivity).

Data scientists call the data used to predict these labels “features.” The most valuable fields (features) identified by the model come from the master data. Therefore, trained algorithms can predict the deviations very early in the

[4] It was a big advantage that the POC could use the ICS rules as labels.

[5] The impairment system calculates the impairment value by using credit risk methods like life-time expected loss for stage 2 and 3 transactions.

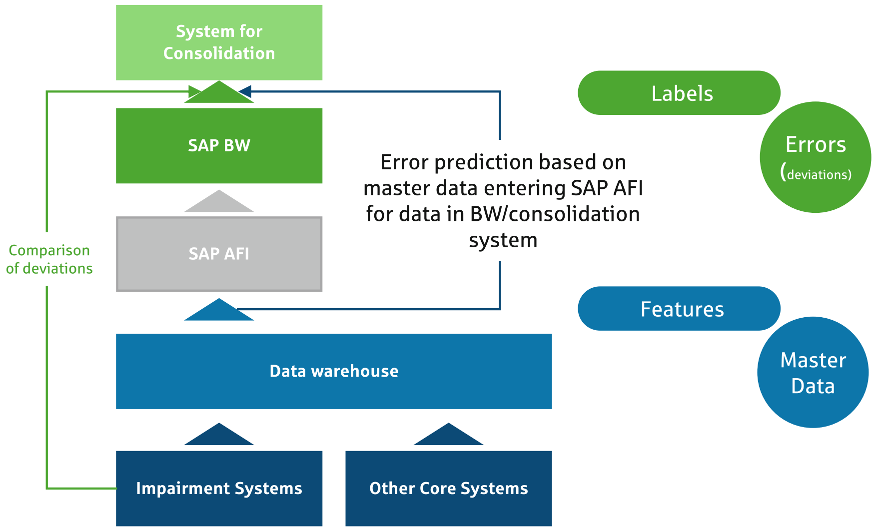

Fig. 5 Systems involved and data flow (© Helaba)

process (based on the core-system data/master data). The algorithm recognizes patterns in the source system data that lead to incorrect changes in the impairment key figures during processing in SAP AFI.

3.2 Data Architecture—Modeling

Given the nature of the POC, the technology sophistication used was pragmatic. As Helaba did not heavily invest in data lake technology at the start of the POC and the POC also restricted the data period to a statistical minimum, the task could be carried out with standard technology.

In the first steps of the POC, the raw data was only provided by text files. During the POC, the raw data was stored in a traditional SQL database (Microsoft SQL server 2014). The raw data was connected by standard SQL statements in the server and some links were carried out as preprocessing to the model in R using the famous package dplyr (Fig. 6).[6]

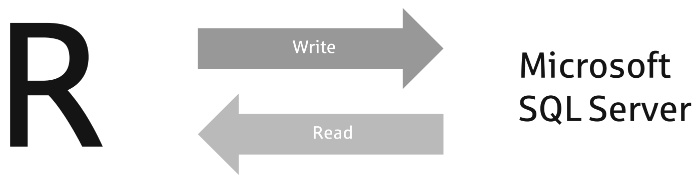

Fig. 6 Tool data and analysis (© Helaba)

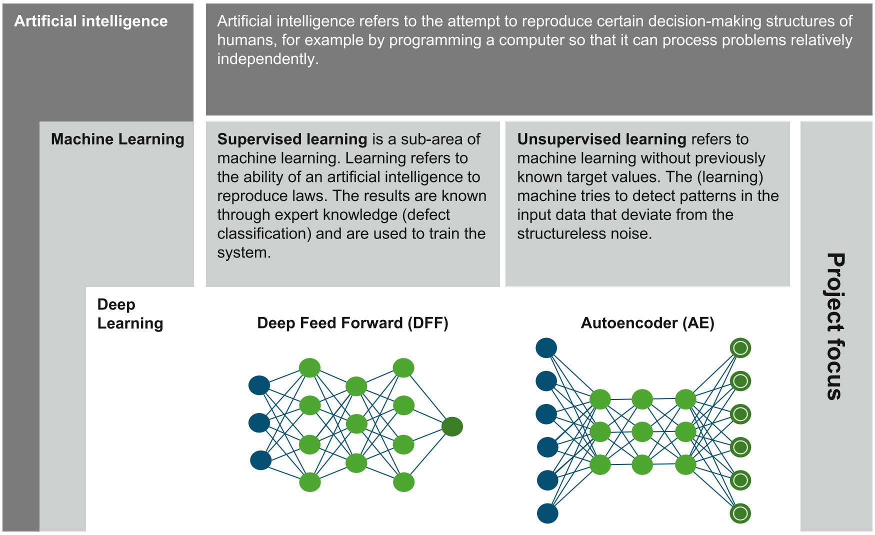

Fig. 7 Overview artificial intelligence (© Helaba)

The data to train the model is provided by the SQL views and some data transformation in R so that the common models (FFN[7] random forest,[8] etc.) can be used. In addition, H20[9] was used for some of the more advanced models. After the cross validation of the model,[10] the model parameters (and hyperparameters) are stored in a text file so the application (error pattern recognition) can be performed using the information in the text file.[11]

3.3 Modeling

In the machine learning environment, there is a wide range of different models available (by academia and by frameworks see Liermann, Overview Machine Learning and Deep Learning Frameworks [2021]). A common structure to differentiate the model is shown on the left side of Fig. 7. Artificial intelligence is the umbrella term, followed by machine learning

[6] See information about the R package (Wickham et al. 2020) and, for an introduction, see Institute for Statistics and Mathematics of WU (2020).

[7] Deep Feed Forward network—see Section 3.1 in Liermann et al., Deep Learning—An Introduction (2019a).

[8] See Section 3.7.1 in Liermann et al., Introduction in Machine Learning (2019b).

[9] H2o.ai is a machine learning framework (see H2O.ai 2019).

[10] See Section 4.1 in Liermann et al., Introduction in Machine Learning (2019b).

[11] It was beyond the scope of the POC to establish a proper infrastructure to organize the parameter and hyperparameter storage and reading.

as a sub-category and further broken down into deep learning (for more details see Section 1.1 in Liermann et al., Introduction in Machine Learning [2019b]).

Both machine learning and deep learning can be separated into supervised learning (e.g., for classification problems a label is given for the classes) and unsupervised learning (e.g., for classification problems no label is defined but the model groups the different elements).

In the POC, supervised learning (classification models) is used to identify the severity of an error in the impairment reporting. In addition, some unsupervised learning methods (classification models, like autoencoder) are applied. Although anomaly detection is a traditional field for autoencoders[12] the performance of the supervised models was far better.

In the outlook (Sect. 6), we will discuss some reasonable applications for unsupervised learning in the context of the impairment process.

4 Project Approach

The journey of the digital transformation of a bank is not easy and not always entirely successful. We see the best efficiency when the journey starts with the business divisions because the value needs to be created there and not only in the IT department. We find it helpful to support the business departments with a staff division focusing on digitalization and helping the business divisions to understand the technology and the contexts in which the tools work best.[13]

The Proof of Concept was embedded within a larger initiative in which the bank sought relevant and fitting use cases to explore new technologies in a practical environment.

The project setup was not purely agile (Scrum, see Akhgarnush et al. [2021]), but we used agile components, such as daily standups and an iterative and Sprint-oriented temporal structure for data discovery, data cleansing and model development.

Figure 8 shows the five steps of a classic machine learning project.

[12] See Section 3.3 in Liermann et al., Deep Learning—An Introduction (2019a).

[13] The opposite approach would be to start with the technology and incorporate the business divisions once the infrastructure is up and running.

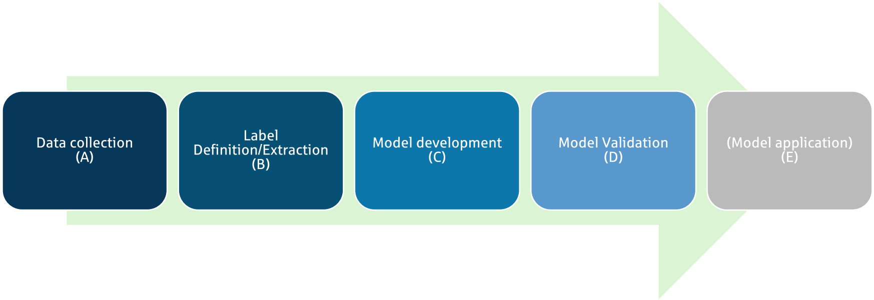

Fig. 8 Project steps (© Helaba)

5 Results and Project Experience

The result of the POC has many dimensions. The major result is that there is potential to discover patterns in data flows. The potential can be realized and leads to a reduction in process time[14] and in manual effort.[15] In addition, the quality of an error prediction[16] has improved. This quality improvement even offers the option to predict the key figure[17] causing the error.

5.1 Modeling Results

The data scientist expresses the quality of a model[18] by the rate of incorrectly predicted values. These incorrectly predicted values are differentiated into (A) if the value should be true (in our case, the transaction has a deviation) but is predicted as false (the model overlooked the error) and (B) the value should be false (the transaction is correct) but the model predicts a true one (the model saw an error but there was none). (A) is called false positive (FP) and (B) is called false negative (FN).

To transfer the absolute figures of the false positive and the false negative into relative figures, the data scientist uses precision (for expressing the false positive) and recall (for expressing the false negative). On the right side of Fig. 9, the model accuracy is shown for two of the observed labels.[19] On the left side of Fig. 9, the key model accuracy figures are displayed.

[14] Errors are discovered earlier in the process.

[15] More steps in the reconciliation are automated.

[16] A distinction can be made between a minor error (not relevant) and significant deviation.

[17] The total impairment value is composed of up to fifteen different key figures. The deviation is usually connected to one key figure that has not been processed properly.

[18] Especially in classification models we applied in the POC.

[19] Even if sixteen transactions were predicted falsely, one should bear in mind that over 80,000 were properly identified.

Concrete project results

- Model Label_2

Deviation of actual and expected values between SAP BW and previous systems

Average significance of the model (December 2018 to September 2019)		
AUC	*Recall*	*Precision*
96.44%	97.66%	86.71%

- Model Label_4

Entries for provisions or value adjustments for risk provisioning on several impairment levels in SAP BW

Average informative value of the model (April to July)		
AUC	*Recall*	*Precision*
99.49%	100%	84.06%

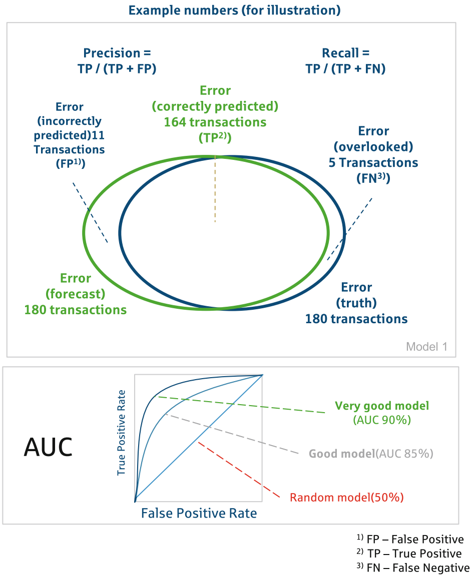

Fig. 9 Modeling results (© Helaba)

The accuracy of the trained models is high. The high accuracy is to keep in balance while decreasing overfitting.[20] It was another modeling challenge to deal with the low quantity of errors (80,000 transactions and approximately 170 significant deviations, corresponding with a rate of 0.2%). However, the data scientist found model types to handle this challenge. This is the art of the data scientist. This aspect will be given more attention in the productive implementation.

5.2 Lessons Learned

As the POC was a complete success,[21] we feel encouraged by the path we took. Starting with the business requirements and pain points is the key to success. Technology is an important component overall, but it has less importance without a business context to show the possibilities of the new technologies.

Given the relatively small data size (especially in the context of a POC), the technology was handled perfectly with standard data storage and open source

[20] Overfitting can occur when the historical data is too similar over time and the model can only predict these patterns but is unable to follow changing patterns.

[21] In 2020, the project started to place the results and processes in a productive environment.

components. It was key to have colleagues with technological skill sets[22] to understand the components used (SQL database and R within Microsoft visual studio). In the long-term outlook, the setup could be improved efficiently by data lake technology and extended infrastructure for data scientists (like Cloudera or other providers of data scientist-related infrastructure).

Open source offers a wide variety of models and tools to get the data processing and model development in place.

If we analyze where we spent most of our time and human resources during the project, it was definitely data and data infrastructure, and we expect this to remain the precondition for good models. However, the models offer sufficient leverage to achieve the business requirements.

6 Summary and Outlook

There are different roads to a successful digital transformation. We are quite happy with the track we took by starting in the business domains. Naturally, a technology-focused approach has advantages too,[23] but we achieved our business needs with minimum technological overheads by using standard technology that is available at the bank as well as open source technology that is approved by the IT department.

The reasonable investment gave us the opportunity to evaluate the right processes and the right technology to improve these processes.

Literature

Akhgarnush, Eljar, Fabian Bruse, and Ben Hofer. 2021. "New Project Structure." In *The Digital Journey of Banking and Insurance, Volume I—Disruption and DNA*, edited by Volker Liermann and Claus Stegmann. New York: Palgrave Macmillan.

H2O.ai. 2019. *h2o.ai Overview*, January 29. Accessed January 29, 2019. http://docs.h2o.ai/h2o/latest-stable/h2o-docs/index.html.

Institute for Statistics and Mathematics of WU. 2020. "The Comprehensive R Archive Network." *dplyr: A Grammar of Data Manipulation.* Accessed October 15, 2020. https://CRAN.R-project.org/package=dplyr.

Liermann, Volker. 2021. "Overview Machine Learning and Deep Learning Frameworks." In *The Digital Journey of Banking and Insurance, Volume III—Data*

[22] The employees involved came from a physics and mathematical background.

[23] Better up and running infrastructure.

Storage, Processing, and Analysis, edited by Volker Liermann and Claus Stegmann. New York: Palgrave Macmillan.

Liermann, Volker, Sangmeng Li, and Norbert Schaudinnus. 2019a. "Deep Learning—An Introduction." In *The impact of Digital Transformation and Fintech on the Finance Professional*, edited by Volker Liermann and Claus Stegmann. New York: Palgrave Macmillan.

———. 2019b. "Introduction in Machine Learning." In *The Impact of Digital Transformation and Fintech on the Finance Professional*, edited by Volker Liermann and Claus Stegmann. New York: Palgrave Macmillan.

Wickham, Hadley, Romain François, Lionel Henry, and Kirill Müller. 2020. "Tidyverse." *dplyr.* Accessed October 15, 2020. https://dplyr.tidyverse.org/articles/dplyr.html.