efl insights 02|2024
An efl – the Data Science Institute Publication

efl the Data Science Institute

The Role of Data and Artificial Intelligence for Progressive Public Affairs

Rethinking Table Retrieval From Data Lakes

Customer Loyalty Trajectory for B2B Firms: Unveiling Patterns of Retention

Finding the Right AI Approach for the German Savings Banks

GOETHE UNIVERSITÄT FRANKFURT AM MAIN
TECHNISCHE UNIVERSITÄT DARMSTADT

Impressum

Redaktion
Prof. Dr. Peter Gomber
Julian Schmidt, M.Sc.

Herausgeber
Prof. Dr. Peter Gomber
Stellvertretender Vorstandsvorsitzender
efl – the Data Science Institute e.V.

Kontakt
info@eflab.de
www.eflab.de

Gestaltung
Novensis Communication GmbH
Weilburg

2. Ausgabe, 2024
Auflage: 200 Stück (Print) / 2.000 Stück (E-Paper)

Printed in Germany
ISSN 1866-1238

Editorial

The Role of Data and Artificial Intelligence for Progressive Public Affairs

Stefan Mai

Dr. Stefan Mai
Head of Executive Board Office,
Head of Public Affairs
Union Asset Management Holding AG

The significance of data and Artificial Intelligence (AI) has a profound impact on all industries, presenting both challenges and opportunities. Given its power and relevance, AI has not gone unnoticed in the public affairs sector. The upcoming German federal election in 2025 brings discussions about AI to the forefront, raising questions about the extent to which data will drive the public affairs field and how it will be handled.

AI-generated data provides valuable insights that can help to identify and model trends, patterns, and preferences among individuals. It enables a deeper understanding of the external environment and allows for more targeted communication and tailored messaging. This can be an efficient way to engage with stakeholders and ensure that information reaches the intended audience effectively. AI can also help to allocate resources of public affairs teams, eventually devoting more time to strategic activities and tactical considerations. However, there is a clear need to address the legal framework for data collection and usage, particularly in terms of data protection.

In contrast to Germany, the United States of America place less emphasis on data protection, resulting in a more prevalent data-based political work. As a consequence, specific interest groups influence the political sphere more than the general American society can do. These circumstances are not desirable for us: We have the responsibility to maintain a perspective on issues relevant to society as a whole and to convey these across the social breadth. Professionals need to resist the temptation to only use AI data to develop strongly individualized content for specific groups. As an interface between politics, business, and society, professionals have the task of continuing to communicate issues in a holistic matter and being aware of its influential power.

Excessive tailoring or adapting of content using AI-generated information poses two risks. Firstly, if lobbying is solely based on available data about individuals, it could impact democracy negatively. Over-reliance on such data might cultivate narrow echo chambers, an environment, where individuals are only exposed to information and perspectives that align with their existing beliefs. These circumstances can reinforce biases and further divide society, hindering the exchange of diverse opinions and limiting democratic discourse. Secondly, the sole consideration of AI data alone might perpetuate patterns and narratives that society aims to overcome, such as marginalization. Especially, since limited or biased data used to train AI models can further preserve bias in AI systems. It is therefore important to acknowledge the fact that AI is not a neutral technology. To use AI effectively and for the public interest, ethical and social implications of its use have to be understood and implemented into a company's AI governance system to ensure accountability, fairness, and transparency in its applications.

Given the listed facts, effective and progressive public affairs work must find a balance of AI-generated data and a network-oriented approach with personal interaction. The face-to-face interaction with stakeholders will still be key in building trust, foster relationships, and establish a deeper understanding of the concerns, needs, and aspirations. It provides an opportunity for listening, empathy, and reciprocal communication. Actively seeking dialogue helps in uncovering unspoken needs, identifying personal feelings and power dynamics, recognizing shared values, and forging alliances or partnerships.

We as public affairs professionals therefore aim to consider the interests of all involved and affected parties, in line with the values and principles of Union Investment: partnership, transparency, solidarity, and down-to-earthiness. These guidelines apply equally in both, the digital and analog worlds, to guarantee a progressive and successful effort in our public affairs work.

Research Report

Rethinking Table Retrieval From Data Lakes

EXISTING TABLE RETRIEVAL APPROACHES ESTIMATE EACH TABLE'S RELEVANCE FOR A PARTICULAR INFORMATION NEED AND RETURN A RANKING OF THE MOST RELEVANT TABLES. THIS APPROACH IS NOT IDEAL SINCE THE RETURNED TABLES OFTEN INCLUDE IRRELEVANT DATA AND THE REQUIRED INFORMATION MAY BE SCATTERED ACROSS MULTIPLE TABLES. TO ADDRESS THESE ISSUES, WE PROPOSE THE IDEA OF FINE-GRAINED STRUCTURED TABLE RETRIEVAL AND PRESENT OUR VISION OF R2D2, A SYSTEM WHICH SLICES TABLES INTO SMALL TILES THAT ARE LATER COMPOSED INTO A STRUCTURED RESULT THAT IS TAILORED TO THE USER-PROVIDED INFORMATION NEED. AN INITIAL EVALUATION OF OUR APPROACH DEMONSTRATES HOW OUR IDEA CAN IMPROVE TABLE RETRIEVAL AND RELEVANT DOWNSTREAM TASKS SUCH AS TABLE QUESTION ANSWERING.

Jan-Micha Bodensohn Carsten Binnig

Introduction
Table retrieval from data lakes has recently become important for many downstream tasks. For example, data engineering efforts often involve retrieving relevant tables from data lakes to provide the basis for further data analysis and exploration. Moreover, table retrieval has also gained importance beyond data engineering. For example, recent open-domain table question answering (TQA) approaches combine table retrieval with large language models (LLMs) to answer questions based on tables.

Existing table retrieval approaches often follow the same procedure: Given a data lake of tables and a user-provided information need, they estimate each table's relevance to the information need and return a ranking of the most relevant tables. While this approach helps users to search through large collections of tables, it still has severe limitations:

(1) *Irrelevant data in full tables.* A major issue of existing table retrieval approaches is that they are coarse-grained and often only return full tables. While this approach is somewhat tractable for small web tables, retrieving full tables can cause high manual filtering overheads for real-world table repositories such as enterprise data lakes, where tables often contain hundreds of columns and millions of rows.

(2) *Scattered information.* A second major issue is that relevant information is often scattered across multiple tables in the ranked result list, causing additional difficulties and overheads for users to identify and combine the relevant information.

Towards Fine-Grained Structured Table Retrieval
To address these limitations of existing table retrieval approaches, we propose the idea of fine-grained structured table retrieval, where the goal is to use data from the data lake tables to construct small relational databases that are tailored to the specific user-provided information needs. We further present our vision of R2D2 (Retrieving Relevant Data from Data Lakes), a system which tackles this goal by diverging from existing table retrieval approaches in two key aspects:

(1) Instead of retrieving full tables, R2D2 first splits the tables into smaller tiles (sub-tables), allowing the retrieval step to select which parts of the table to return to the user. In contrast to existing table segmentation approaches, we propose a novel representation-based slicing algorithm that is agnostic to the information needs but takes the table data into account to group related rows and columns in the same tiles.

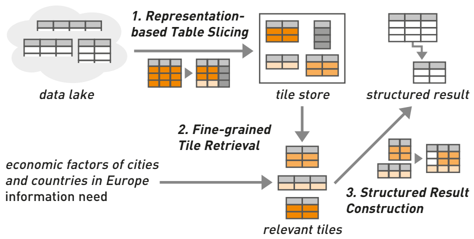

Figure 1: Prototype Overview

(2) Instead of returning ranked lists of independent tables as a result, R2D2 detects relationships between the individual tiles in the result and uses them to compose structured results in the form of small relational databases. Returning multiple tables allows us to better capture the relevant entities of the information needs. Moreover, we argue that such structured results come more naturally to users and let them get an overview of the retrieved data more easily.

Prototype Implementation
As shown in Figure 1, R2D2 works in three stages. The first stage uses our novel representation-based slicing algorithm to slice each table into a set of small tiles. The second stage then uses embeddings of the information need (i.e., the user's question) and tiles to retrieve a set of relevant tiles. Finally, the third stage derives the relationships between the retrieved tiles to compose the structured result that represents the data satisfying the user's question.

Representation-based slicing. By slicing the tables into tiles, we allow the retrieval step to filter which parts of the table to return to the user. Naively partitioning the tables into row-wise or column-wise chunks may leave related information, that is relevant to the same information needs, scattered across multiple tiles. To tackle this issue, we propose a novel slicing algorithm that works based on representations of the table data. In each step, we partition the table either horizontally or vertically to incrementally increase the focus of the resulting tiles whilst keeping related rows and columns in the same tile.

The slicing algorithm starts with a full table and then recursively splits it into multiple tiles. In each step, we first determine which tile to split. Since we want to keep related information in the same tile, we compute a measure of semantic heterogeneity d^i for each tile and split the tile with the highest heterogeneity. To operationalize this, we compute representations for each tile's rows and columns using a pre-trained language model (Reimers and Gurevych, 2019). Next, we determine the maximum pairwise cosine distances d^i_{row} and d^i_{col} between the row and column representations of tile i and consider the greater of the two distances as the tile's semantic heterogeneity d^i. We also use d^i_{row} and d^i_{col} to determine whether to split the tile row-wise or column-wise based on which of the two values is greater. To then actually split the tile, we cluster the tile into k clusters of rows/columns based on the tile's row/column representations. We then continue to recursively split the resulting tiles until a specified number of tiles is reached.

Fine-grained tile retrieval. Given a user-provided information need, R2D2 performs a cosine similarity-based vector search to retrieve a set of relevant tiles. To compute the representation for the information need, we simply embed the natural language string. To compute the tile representations, we compute embeddings for all rows and columns of the source table. The tile representations are then computed by averaging the embeddings of all corresponding rows and columns.

Structured result construction. Finally, R2D2 uses the retrieved tiles to compose the structured result that is returned to the user. The main idea is that instead of returning a ranked list of individual tiles, we aim to construct a structured set of tables that better represents the relationships between the data. Our prototype of R2D2 merges all tiles that come from the same source table and thus have explicit structural relationships. For each source table, we determine the smallest sets of rows and columns so that their overlap contains all retrieved tiles and construct one result table based on this overlap.

Table Question Answering

To demonstrate the viability of the ideas behind our approach, we combine R2D2 with a LLM to evaluate how our approach can benefit TQA as a first downstream task.

Data sets. We experiment on two data sets: Spider (Yu et al., 2018) is a text-to-SQL data set containing relational databases with natural language questions and matching SQL queries. Open-WikiTable (Kweon et al., 2023) is an open-domain TQA data set containing web tables with natural language questions and matching SQL queries.

Experimental setup. We combine our table retriever (R2D2) with a LLM (OpenAI's GPT-4 Turbo) in a retriever-reader framework, i.e., the LLM receives the retrieved tables and the question and generates the answer. To analyze the effectiveness of the retrieval approach, we constrain the input size to 1,000 tokens for Open-WikiTable and 200 tokens for Spider. We compute the exact match (EM) accuracy to evaluate the TQA and measure the recall (R) based on which cells are included in the model input.

Results. Table 1 shows that R2D2 vastly improves the accuracy and recall on both data sets compared to retrieving full tables (no slicing). Furthermore, we see that our representation-based slicing algorithm outperforms naïve row-wise and column-wise slicing on Open-WikiTable. On Spider, our approach outperforms row-wise slicing and performs on par with column-wise slicing. Finally, we see that omitting the result construction and instead using individual tiles as input to the LLM leads to sharp declines in accuracy and recall, indicating the importance of the result construction step.

	OWT		Spider	
	EM	**R**	**EM**	**R**
No Slicing	.08	.07	.08	.44
Row-Wise Slicing	.29	.38	.09	.64
Column-Wise Slicing	.25	.48	**.13**	**.84**
R2D2 W/O Result Construction	.12	.21	.11	.65
R2D2	**.32**	**.51**	**.13**	.82

Table 1: TQA EM Accuracy and Recall

References

Kweon, S.; Kwon, Y.; Cho, S.; Jo, Y.; Choi, E.:
Open-WikiTable: Dataset for Open Domain Question Answering with Complex Reasoning over Table.
Working Paper, 2023.

Reimers, N.; Gurevych, I.:
Sentence-BERT: Sentence Embeddings Using Siamese BERT-Networks.
In: Proceedings of the 14th Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP); Hong Kong, China, 2019.

Yu, T.; Zhang, R.; Yang, K.; Yasunaga, M.; Wang, D.; et al.:
Spider: A Large-Scale Human-Labeled Dataset for Complex and Cross-Domain Semantic Parsing and Text-to-SQL Task.
In: Proceedings of the 23rd Conference on Empirical Methods in Natural Language Processing (EMNLP); Brussels, Belgium, 2018.

Research Report

Customer Loyalty Trajectory for B2B Firms: Unveiling Patterns of Retention

CUSTOMER LOYALTY IS A CRITICAL MEASURE FOR SUCCESS, SHOWING IF A FIRM'S PRODUCT IS RECEIVED WELL BY ITS CUSTOMERS. TO UNDERSTAND ITS DEVELOPMENT OVER TIME, TWO FUNDAMENTAL QUESTIONS MUST BE ANSWERED: (I) HOW WILL CURRENT CUSTOMERS' LOYALTY DEVELOP, AND (II) WILL NEW CUSTOMERS' LOYALTY DIFFER FROM CURRENT CUSTOMERS' LOYALTY? THE AUTHORS EMPIRICALLY ANSWER THESE QUESTIONS BASED ON A DATA SET INCLUDING ~500 B2B WEB TECHNOLOGIES WITH JOINTLY ~325 MILLION CUSTOMERS SPANNING OVER 24 YEARS. THEY SHOW THAT LOYALTY HARDLY DEVELOPS AND, IF SO, IT RATHER DECREASES THAN INCREASES. THE LOYALTY OF CURRENT CUSTOMERS RARELY CHANGES AND, IF SO, RATHER INCREASES THAN DECREASES. NEW CUSTOMERS ARE MOST LIKELY LESS LOYAL THAN CURRENT CUSTOMERS. THESE RESULTS SHOW THAT BY FAILING TO ACCOUNT FOR THESE UNDERLYING DEVELOPMENTS, STAKEHOLDERS, IN MOST CASES, DRAW THE WRONG CONCLUSIONS ABOUT PRODUCT VALUE MEASURED VIA CUSTOMER LIFETIME VALUE.

Johannes Roscher

Ali Tamaddoni

Bernd Skiera

Introduction

The success of a firm depends on its customers. The development of the number of customers is therefore highly relevant for managers, investors, and analysts. While firms can often influence the number of newly acquired customers by the size of the marketing spend, observing customer loyalty provides insights into the underlying trajectory (Ellis and Brown, 2017; Ries, 2011; TechCrunch, 2022). Observing customer loyalty across customers acquired at different points in time can, however, lead to the wrong conclusions if its underlying forces are overlooked (McCarthy et al., 2024).

Suppose an investor would, for example, look at the loyalty development of Company A in period 5 in Figure 1. In that case, he/she observes a customer loyalty rate that seems constant at ~80% in the first five periods. However, the observed loyalty can decrease significantly in the following years. Why? The overall loyalty development is driven (I) by an increase in loyalty of current customers and (II) by a decrease in loyalty of new customers compared to current customers. While the contrary development of the two drivers equals out at the beginning of the firm's lifecycle, over time, the decrease in new customers' loyalty outweighs the gain in current customers' loyalty, leading to an overall lower loyalty rate. The contrary loyalty dynamics mislead the stakeholder to an over- or underestimation dependent on the measurement method used for firm success.

A priori, sound reasoning exists for each direction of the loyalty movement. Current customers' loyalty could increase over time due to survival bias, with customers with a lower product fit churning early on and customers with a higher fit remaining (Fader and Hardie, 2007; Gupta and Lehmann, 2005; Kumar and Reinartz, 2018; McCarthy et al., 2024). However, on the contrary, customer saturation could increase with increasing duration of customer affiliation, and the product (which could be a good software or service) becomes less valuable, leading to decreasing customer loyalty (McCarthy et al., 2024).

For new customers' loyalty development, loyalty could increase as the product first must be

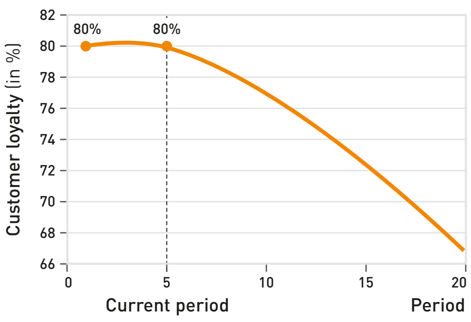

Figure 1: Exemplary Customer Loyalty Development of Company A

established in the market (Prins et al., 2009) before attracting more established customers with higher loyalty. Otherwise, loyalty could decrease as the customers with the highest fit become customers early on, with later acquired customers being less loyal (Eisenmann, 2021).

So, insights into which direction the loyalty of current and new customers develops help investors, analysts, and managers to better evaluate future firm performance. Therefore, we aim to answer the following research questions:

1. How will current customers' loyalty develop over time?
2. How will new customers' loyalty differ from current customers' loyalty?
3. Which drivers impact loyalty development? (e.g., technology types, customer growth, location of technology's headquarters)
4. How will the change in current and new customers' loyalty impact overall loyalty?

Research Approach

To answer these questions, we construct period-cohort tables, whereby we group customers into cohorts based on the year of their first purchase. They are then tracked along their age (i.e., the time they have been customers), showing the number of customers who remained active in the next year out of the initially acquired customers of the cohort. We convert these tables to year-to-year retention tables, showing what proportion of customers of a cohort in a period remained active in the next period (note, the value in each cell is between 0 if all customers churn, and 1 if all customers stay).

The development of retention rates along the table is equal to the development of current customers' loyalty (hereafter the age effect – corresponding to research question 1), while the difference in retention rate between the cohorts is equal to the difference in new customers' loyalty compared to current customers' loyalty (hereafter the cohort effect – corresponding to research question 2).

Data Set

To empirically answer our research questions, we assemble a B2B data set including 689 web technologies, hereafter referred to as technologies each treated as an independent entity, out of which we keep (after several data cleaning steps) 497 technologies with more than 325 million customers. The data set comprises eight technology types (e.g., web hosting and infrastructure, advertising and marketing, content management and development) and spans 24 years (2000 to 2023).

Looking at the average churn rates across technologies, we see that after 3 years, on average, 55% of customers churned. Looking at average customer loyalty by technology type, we see only small differences between categories: All average loyalty rates are between 73% and 75%, apart from Email Hosting at 90% and Media and Content Delivery at 64%. In addition, we can observe a difference between the average loyalty rate of technologies in the US (76%) and other geographic areas (80%), which is significant with a p-value of 0.04% based on the Mann-Whitney U test.

Results

To assess the direction of the age and cohort effect across our full sample, we run individual regressions for each technology, summarizing the direction of effects in Figure 2. It shows the number of significantly positive, neutral, and significantly negative coefficients for the age and cohort variable of the individual models across our sample. First, looking at the age and cohort effect separately, we see that 279 of the 497 technologies have a neutral, i.e., no age effect. Out of the remaining 218 technologies, 162 show a positive and 56 a negative age effect.

We conclude that the age effect is primarily neutral but more likely positive than negative. The cohort effect, on the other hand, is negative for 260 technologies and neutral for 203 technologies. Looking at the interaction of the effects, we see that if the cohort effect is positive, the age effect is also often positive. When the cohort effect is neutral, the age effect tends to be neutral or positive. In contrast, if the cohort effect is negative, the age effect is reasonably balanced between positive and negative; however, it is mostly neutral.

To analyze the impact of potential drivers on the age and cohort effect, we run three separate regressions with all drivers as independent variables of each regression and the age, cohort, and constant of the previously derived age-cohort OLS regression models as the dependent variables. As drivers, we investigate (I) the year of technology establishment, (II) the technology origin country, (III) the growth of the company, (IV) the age of customers, and (V) the technology type.

We show that the year a technology has been established, the technology's origin country, and the average age of cohorts' customers only drive overall loyalty. Technologies established later have lower loyalty, technologies with their firms' headquarters in the US have lower loyalty, and technologies with a higher average age of

N = 497

		Age effect			
		Positive	**Neutral**	**Negative**	**Total**
Cohort effect	**Positive**	27	6	1	34
	Neutral	81	115	7	203
	Negative	54	158	48	260
	Total	162	279	56	497

Number of observations: Low High

Figure 2: Overview of the Direction of Age and Cohort Effect (Significance Level >5% and Positive/Negative Coefficient >0.3% Required for Non-neutral Classification)

N = 497

		Cohort effect	
	Technologies	**Not considered**	**Considered**
Age effect	**Not considered**	***Constant loyalty assumption***	322 *(65%)* 175 *(35%)*
	Considered	200 *(40%)* 297 *(60%)*	188 *(38%)* 309 *(62%)*

Limited effect (+/-10% deviation from constant loyalty development)

Major effect (> +/-10% deviation from constant loyalty development)

(x%) Share of technologies of total

Figure 3: Overview of the Number of Technologies by Deviation from CLV, Including Age and Cohort Effects (Assuming Time Shift for Cohort Effect of 5 Periods, Maximum Retention Rate of 99%, and a Discount Rate of 10%)

customers (i.e., how long the customer existed before being acquired) have a higher loyalty. The type of technology, however, influences the cohort effect and the overall loyalty, with the direction and significance of the effect depending on the specific technology type. The growth of the company has no significant impact on the age effect, cohort effect, or overall loyalty level. The age effect is not significantly impacted by any of the investigated drivers overall.

To see whether the age or cohort effect has a stronger impact on the overall retention rate across technologies, we run an OLS regression with the overall retention development analysis coefficients as dependent and the age and cohort effect coefficients as independent variables. We see that the age and cohort coefficients are both highly significant, with the cohort coefficient having a higher z-value (15.686 vis-à-vis 4.099) and a higher coefficient (0.704 vis-à-vis 0.1805). We conclude that the cohort effect is more relevant for the overall retention development than the age effect in our sample. This conclusion corresponds with our observation of a higher share of technologies with a negative overall loyalty trend, driven by the negative cohort effect.

Managerial Implications

We show that the age effect tends to be neutral or positive, and the cohort effect tends to be neutral or negative, with the overall loyalty being more likely to develop negatively than positively but keeping a neutral development in most cases. The question, therefore, is: Do stakeholders even need to care about the effects or do the opposing age and cohort effects counterbalance each other?

To investigate this question, we calculate four exemplary customer lifetime values (CLVs) for each technology, by (I) not including any age or cohort effects, (II) including only the respective age effect, (III) including only the respective cohort effect, and (IV) including both effects. We assume the base loyalty level, the age, and the cohort effect to correspond to the intercept, the age, and the cohort coefficient of our initial age-cohort OLS regression per technology. After calculating the four CLVs per technology, we calculate the differences for the CLVs, including only the age effect, the cohort effect, and the age and cohort effect compared to the CLV, excluding both effects.

In Figure 3, we show, dependent on the included effects, the share of technologies for which the deviation in CLV compared to the constant loyalty CLV is larger than 10%. Accordingly, if only considering the age effect, we see that the CLV deviates for 60% of technologies by more than 10%. The standalone cohort effect leads to a deviation of >10% for 35% of technologies. If both effects are included, we see that the CLV is significantly impacted for 62% of technologies. We observe that 33% of technologies deviate by ±10-25% compared to a constant loyalty assumption, and 29% deviate by more than 25%. In addition, we see an equal distribution in the number of technologies between technologies deviating by more than +10% (32%) and by deviating more than -10% (31%).

We conclude that the age and cohort effect significantly impact the CLV calculation in most cases, highlighting the need for stakeholders to consider them in their assessments. Startup investors, in particular, should request customer loyalty data not only on an overall level but also cohort-by-cohort to predict future development. As shown in the initial example of this paper, a contrary age and cohort effect can, jointly with customer growth, lead to the wrong conclusions, e.g., by assuming a firm has constant customer loyalty when, in reality, the customers with the highest product-fit have already been acquired, leading subsequently to customers with lower loyalty. We, therefore, argue that outside stakeholders should urge firms to publish their customer loyalty developments on a cohort level, providing insights into the underlying dynamics of the firm trajectory.

References

Eisenmann, T.:
Why Startups Fail: A New Roadmap for Entrepreneurial Success.
Crown Currency, New York (NY), 2021.

Ellis, S.; Brown, M.:
Hacking Growth.
Virgin Books, London, 2017.

Fader, P. S.; Hardie, B. G. S.:
How to Project Customer Retention.
In: Journal of Interactive Marketing, 21 (2007) 1, pp. 76–90.

Gupta, S.; Lehmann, D.:
Managing Customers as Investments: The Strategic Value of Customers in the Long Run.
FT Press, New Jersey (NJ), 2005.

Kumar, V.; Reinartz, W.:
Customer Relationship Management.
Springer, Berlin, 2018.

TechCrunch:
How to Improve Retention, Growth Marketing's Golden Metric, https://techcrunch.com/2022/06/02/how-to-improve-retention-growth-marketings-golden-metric, 2022.

McCarthy, D. M.; Libai, B.; Schoenmueller, V.:
The Declining Churn Fallacy over the Product Lifecycle.
Working Paper, 2024.

Prins, R.; Verhoef, P. C.; Franses, P. H.:
The Impact of Adoption Timing on New Service Usage and Early Disadoption.
In: International Journal of Research in Marketing, 26 (2009) 4, pp. 304–313.

Ries, E.:
The Lean Startup: How Today's Entrepreneurs Use Continuous Innovation to Create Radically Successful Businesses.
Crown Currency, New York (NY), 2011.

Insideview

Finding the Right AI Approach for the German Savings Banks

INTERVIEW WITH MARTIN WALDMANN

Martin Waldmann
Executive Director
Finanz Informatik

Generative AI is a game changer – also in the financial sector. Institutions and their IT service providers need to consider carefully: Which AI approach will enable them to implement optimal solutions for themselves and their customers in this highly regulated environment?

How did Finanz Informatik, as the savings banks' digitalization partner, proceed here?

As the savings banks' digitalization partner, we aim to digitalize processes and workflows in the savings banks, which are currently only partially automatized, in a customer-centric and practice-oriented manner. Generative AI offers new opportunities, whereby deep integration and access to data are crucial for increasing efficiency and offering new solutions with AI.

Can this be achieved with AI solutions from the cloud?

Today's general-purpose AI models are focused on the consumer market and are connected to a cloud infrastructure. The range of services provided is enormously. For the savings banks, we need more specific AI models that have been additionally trained with savings bank data and are tailored to their needs.

How did you proceed here?

As the savings banks' digitalization partner, we already supply all banking solutions centrally today – from design and development to operations. As a result, we have established a market leading expertise and a scalable infrastructure. Our AI strategy for the savings banks builds upon that and relies on powerful pre-trained open source AI models from Meta and Mistral, which we train with the savings banks' data in our own data centers to create a "savings bank AI model".

What are the advantages of this approach?

We can tailor our AI solutions precisely to the needs of the savings banks and their customers. Thanks to the deep integration of the "savings bank AI model" into our IT platform, we can seamlessly integrate AI solutions into the existing and future business processes. This applies to stationary sales as well as digital channels and service centers.

In the next quarter, we will also provide the savings banks' employees with the "S-KIPilot", which supports the classic generative AI functions and can successively access the savings banks' data via interfaces.

Does this approach offer further advantages?

The "S-KIPilot" is based on our specific "savings bank AI model" and is fully operated in our own data centers. This ensures a high level of data protection and information security because no customer or contract data is transferred to any cloud. We also retain full control over the AI's behavior – this allows us to better understand results and reduce "hallucinations" to a minimum. We have currently integrated different open source AI models from two providers in our AI platform which we actuate in a controlled manner. This allows us to reduce the dependency on single providers or AI models.

Where do you see the benefits of generative AI for savings banks in the future?

With our "savings bank AI model" and the deep integration into our IT platform, we can link the savings banks' data and information intelligently. In this, we see outstanding potential, which we want to use to further increase efficiency and customer satisfaction for the savings banks.

Thank you for this interesting conversation.

Infopool

News

EFL ANNUAL CONFERENCE 2024

The efl Annual Conference on "More Value for Money in Financial Advice – How Data Science Can Pave the Way" will take place on November 15th, 2024. The focus of the conference will be on data science in financial consulting, which goes beyond the use of AI. We thank Prof. Dr. Andreas Hackethal and his team for organizing this year's conference. For more information, please visit: **https://www.eflab.de/events/efl-annual-conference-2024**.

Hesse Funds Projects of Prof. Dr. Hinz

The Hessian Loewe program funds the project "TransfAIr – Development and Design of a Software-supported Compliance Check for Transparent and Fair AI Systems" of Prof. Dr. Oliver Hinz with an amount of EUR 690,000. Furthermore, Hessian.AI funds the project "EcOOL – Economics of Optimizing Organizational LLMs" which examines an optimal fine-tuning of large language models for the industry.

Prof. Dr. Binnig Elected to VLDB Board of Trustees

Prof. Dr. Carsten Binnig has been elected to the Board of Trustees of the US-based Very Large Data Base (VLDB) Endowment for a six-year term. The non-profit organization aims to promote scientific work and international exchange in the field of databases and related areas. The foundation serves as the steering committee for the VLDB conference series, one of the most important and international conferences in the field.

8th SAFE Market Microstructure Conference

Together with the Leibniz Institute for Financial Research SAFE and the Institute for Financial Innovation & Technology at Ludwig Maximilian University of Munich, Prof. Dr. Gomber is organizing the 8th SAFE Market Microstructure Conference. The conference will be held on September 20th, 2024, in Munich. The conference aims to present latest research and to stimulate the discussion on current developments in the field of market microstructure.

Paper Cited by Important Media Channels

Bloomberg and FAZ cited the efl Paper "Shifting Volumes to the Close: Consequences for Price Discovery and Market Quality" by Micha Bender, Dr. Benjamin Clapham, and Benedikt Schwemmlein. The authors investigate the massive shift of trading volumes from continuous trading into the closing auction and the effects of these shifts on the efficiency of closing price determination and market quality in continuous trading. The paper is available here: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4757345.

Selected efl Publications

Abdel-Karim, B. M.; Pfeuffer, N.; Carl, K. V.; Hinz, O.:
How AI-Based Systems Can Induce Reflections: The Case of AI-Augmented Diagnostic Work.
In: Management Information Systems Quarterly (MISQ), 47 (2023) 4, pp. 1395–1424.

Bauer, K.; von Zahn, M.; Hinz, O.:
Expl(AI)ned: The Impact of Explainable Artificial Intelligence on Users' Information Processing.
In: Information Systems Research, 34 (2023) 4, pp. 1582–1602.

Bender, M.; Cestonaro, T.; Schmidt, J.:
Lead-Lag Relationships in Market Microstructure.
In: 2023 London/Oxford/Warwick Financial Mathematics Workshop; London, UK, 2023, and 63rd Annual Meeting of the Southwestern Finance Association (SWFA), Las Vegas (NV), United States, 2024.

Hanneke, B.; Chuang, Y.-J.; Skiera, B.; Hinz, O.:
Resale Royalties in the Creator Economy: Evidence from Digital Asset Markets.
In: Proceedings of the 53rd European Marketing Academy Conference (EMAC); Bucharest, Romania, 2024.

Hättasch, B.; Binnig, C.:
Box2Go: Collaborative Interactive Infobox Filling.
In: Proceedings of the 33rd ACM on Web Conference (WWW); Singapore, Singapore, 2023.

Jöntgen, H.; Lingau, N. V.; Hinz, O.; Holten, R.:
This Is Why We Pay – Motivational Factors for Supporting Subscription-Based Crowdfunding Campaigns.
Forthcoming in: Electronic Markets.

Langenecker, S.; Sturm, C.; Schalles, C.; Binnig, C.:
Pythagoras: Semantic Type Detection of Numerical Data in Enterprise Data Lakes.
In: Proceedings of the 27th International Conference on Extending Database Technology (EDBT); Paestum, Italy, 2024.

Proffen, C.; Jürgensmeier, L.:
Do Political Conflicts Influence Daily Consumption Choices? Evidence from US-China Relations.
In: Journal of Economic Behavior & Organization, 220 (2024) 4, pp. 660–674.

Skiera, B.; Jürgensmeier, L.:
Teaching Marketing Analytics: A Pricing Case Study for Quantitative and Substantive Marketing Skills.
Forthcoming in: Journal of Marketing Analytics.

Weiler, M.; Jansen, N.; Hinz, O.:
Can We Measure the Structural Dimension of Social Capital with Digital Footprint Data? – An Assessment of the Convergent Validity of an Indicator Extracted from Digital Footprint Data.
Forthcoming in: Schmalenbach Journal of Business Research.

For a comprehensive list of all efl news postings and efl publications see: *https://www.eflab.de*

Infopool

RESEARCH PAPER:
CYBERSECURITY RISK

Most firms have a high exposure to cybersecurity risks and were directly or indirectly affected by a cyberattack. The authors of this paper examine whether cybersecurity risk of firms is priced in the cross-section of their stock returns. Therefore, they construct a novel firm-level measure of these risk factors using textual analysis of cybersecurity risk disclosures (e.g., 10-K statements). The authors' results indicate that cybersecurity risk is priced in the cross-section of stock returns. The measure exhibits time-series and cross-sectional characteristics, such as a positive trend over time and a more frequent occurrence in information technology industries. In addition, the measure can forecast the likelihood of a company being subjected to a future cyberattack.

Florackis, C.; Louca, C.; Michaely, R.; Weber, M.
In: The Review of Financial Studies, 36 (2023) 1, pp. 351–407.

RESEARCH PAPER:
CROWDSOURCING PEER INFORMATION TO CHANGE SPENDING BEHAVIOR

The authors investigate the behavioral effects of providing consumers with crowdsourced information about the consumption patterns of unknown demographic peers in an app. The results show that app users, who have spent much more on consumption than their peers, reduce consumption after being confronted with peer averages. Thereby, users only adjust behavior if the benchmark is peer-group specific but not if it refers to general population. Within 12 months of information provision, overspenders close 17% and underspenders 5% of their gap relative to peers. This nudge works purely through the information channel and is not driven by confounding events or social norms. The general implication is that app dashboards, which present relevant social benchmark information, can promote individually desirable behavior.

D'Acunto, F.; Rossi, A. G.; Weber, M.
In: Journal of Financial Economics, 157 (2024) 7, Article No. 103858.

efl insights

The efl publishes the insights in the form of a periodic newsletter which appears two times a year. Besides a number of printed copies, the efl insights is distributed digitally via E-mail for reasons of saving natural resources. The main purpose of the efl insights is to provide latest efl research results to our audience. Therefore, the main part is the description of two research results on a managerial level – complemented by an editorial, an interview, and some short news.

For receiving our efl insights regularly via E-mail, please subscribe on our homepage www.eflab.de (> NEWS > SIGN UP EFL INSIGHTS) as we need your E-mail address for sending the efl insights to you. Alternatively, you can mail your business card with the note "efl insights" to the subsequent postal address or send us an E-mail.

Prof. Dr. Peter Gomber
Vice Chairman of the efl – the Data Science Institute
Goethe University Frankfurt
Theodor-W.-Adorno-Platz 4
D-60629 Frankfurt am Main

insights@eflab.de

Further information about the efl is available at
www.eflab.de.

HOUSE OF FINANCE

The efl – the Data Science Institute is a proud member of the House of Finance of Goethe University, Frankfurt.
For more information about the House of Finance, please visit www.hof.uni-frankfurt.de.

THE EFL – THE DATA SCIENCE INSTITUTE IS AN INDUSTRY-ACADEMIC RESEARCH PARTNERSHIP BETWEEN FRANKFURT AND DARMSTADT UNIVERSITIES AND PARTNERS DEUTSCHE BÖRSE GROUP, DZ BANK GROUP, AMERICAN EXPRESS, DEUTSCHE BUNDESBANK, AND FINANZ INFORMATIK LOCATED AT THE HOUSE OF FINANCE, GOETHE UNIVERSITY, FRANKFURT.

For further information please contact:

Prof. Dr. Peter Gomber
Vice Chairman of the
efl – the Data Science Institute
Goethe University Frankfurt
Theodor-W.-Adorno-Platz 4
D-60629 Frankfurt am Main

Phone +49 (0)69 / 798 - 346 82
E-mail gomber@wiwi.uni-frankfurt.de

Press contact
Phone +49 (0)69 / 798 - 346 82
E-mail presse@eflab.de

or visit our website
http://www.eflab.de