

Editorial

Liebe Leserinnen und Leser,

vor Ihnen liegt nunmehr die bereits fünfzehnte Ausgabe des E-Journals **Anwendungen und Konzepte in der Wirtschaftsinformatik (AKWI)**.

Wie schon in den bisherigen Ausgaben werden viele charakteristische Bereiche der Wirtschaftsinformatik behandelt: Geschäftsprozesse im weiteren Sinne, Systeme wie das SAP System, der Betrieb von Systemen und neuere Ansätze aus dem Bereich der künstlichen Intelligenz. Dies erfolgt in verschiedenen Anwendungen wie logistische Fragestellungen und im Rahmen von Forschungsarbeiten sowie Projektarbeiten mit Industrieunternehmen.

Eine Arbeit behandelt eine ereignisdiskrete Simulation von unterschiedlichen Layouts von Hochregallagern zur Erreichung eines möglichst hohen Durchsatzes. In einem Artikel wird für eine konkrete Unternehmenssituation die Frage behandelt, ob Software Microsoft Power BI oder SAP Analytics Cloud die bessere Wahl von Business Intelligence (BI)-Lösungen ist, um kommende Trends zu erkennen und strategische Entscheidungsfindung zu unterstützen. Eine weitere Arbeit berichtet über Erfolgsfaktoren für die Entwicklung, das Design und die Implementierung einer ML-basierten Predictive-Analytics-Lösung zur Orchestrierung von HMC in Entscheidungsprozessen. Zur Nutzung von komplexen Unternehmensdaten unter anderem in Enterprise Resource Planning (ERP) Systemen sind Prozess- und Sensordaten strukturiert zu sammeln. Dies durch geeignete Sensoren zu tun und mittels künstlicher Intelligenz auszuwerten erfolgt im Forschungsprojekts KI-VISOPRO über das in dieser Ausgabe berichtet wird. In einem weiteren Projekt wurde eine Automatisierung im Einkauf bei der OSRAM Continental GmbH durch Robotic Process Automation realisiert. Diese Automatisierung hat gezeigt, dass mittels Robotic Process Automation das Automatisierungspotential eindeutig erweitert werden kann. Jedoch wurde ebenso deutlich, dass die Technologie eindeutig auf bestimmte Aufgaben begrenzt ist und daher ein hybrider Automatisierungsansatz mit mehreren Technologien die beste Lösung darstellt. Gerade in diesen Tagen hat die Reduktion von Energie eine sehr hohe Bedeutung. Über eine Forschungsarbeit sowohl zur Senkung des Kraftstoffverbrauchs als auch zur Reduzierung der Kraftstoffreserve in der Luftfahrt wird berichtet. Der dabei vorgestellte Ansatz erweist sich aufgrund einer umfangreichen empirischen Analyse als sehr wirkungsvoll. Unternehmen wollen die sehr wichtige Aufgabe der Projektkoordination in ein sogenanntes Project Office verlagern. Durch ein Vorgehensmodell sollen vor allem klein- und mittelständische Unternehmen darin unterstützt werden.

Ein Schwerpunkt dieser Ausgabe liegt im Betrieb und Nutzung von Systemen im weiteren Sinne. So wird die Modifizierung und Verbesserung des Softwarelizenzmanagements durch den Einsatz von Managed Services behandelt. Die Konzeption und Implementierung eines flexiblen Lagerdashboards für SAP EWM mit Microsoft Power BI wird vorgestellt. Dem schließt sich die Beschreibung der Implementierung eines Visualisierungskonzepts für die grafische Darstellung von Daten und Informationen innerhalb des Controllings an. Data-Warehouse-Landschaften in der Cloud am Beispiel von Microsoft Azure Synapse Analytics und Snowflake werden vergleichend bewertet. Wie in SAP Fiore eine moderne „State of the Art“ SAP S/4HANA-Cloud-Anwendung mit einer Benutzeroberfläche konzipiert und implementiert werden kann wird erläutert. Oftmals werden flexible Infrastrukturen in analytischen Systemen in Cloud-Umgebungen benötigt. Eine Möglichkeit zu ihrem automatisierten Aufbau wird vorgestellt. Schließlich rundet die Vorstellung der Standardisierung von automatisierten Bestellprozessen mit Peppol am Beispiel der Friedhelm Loh Group diesen Schwerpunkt ab.

Über Ihr Interesse an der Zeitschrift freuen wir uns und wünschen Ihnen Freude bei der Lektüre.

Regensburg, Fulda, Luzern und Wildau, im Dezember 2021.

Frank Herrmann, Norbert Ketterer, Konrad Marfurt und Christian Müller



Christian Müller



Konrad Marfurt



Norbert Ketterer



Frank Herrmann

SYNERGY BETWEEN SHUTTLES AND STACKER CRANES IN DYNAMIC HYBRID PALLET WAREHOUSES: CONTROL STRATEGIES AND PERFORMANCE EVALUATION

Giulia Siciliano, Yue Yu and Johannes Fottner
Chair of Materials Handling, Material Flow, Logistics
Technical University of Munich
Boltzmannstraße 15, Garching bei München, Germany
E-mail: giulia.siciliano@tum.de

KEYWORDS

Dynamic Hybrid Pallet Warehouse, Shuttle, Stacker Crane, Discrete Event Simulation, Control Strategies

ABSTRACT

This article considers two dynamic hybrid pallet warehouses obtained hybridizing a shuttle-based warehouse with stacker cranes. We begin by describing their design and characteristics. Afterwards, we explain the control algorithms that were developed for them. Next, we illustrate the modalities of the discrete event simulation study we ran to investigate their performance. In conclusion, we discuss the results in terms of throughput of the simulation study to individuate the field of application for the two layouts of dynamic hybrid pallet warehouses in comparison to stacker crane-based and shuttle-based warehouses.

INTRODUCTION

A dynamic hybrid pallet warehouse (DHPW) is a new kind of storage and retrieval system that has a shuttle tier on the base connected to the overlying storage layers through satellite stacker cranes. This arrangement allows a combination of the advantages of shuttle-based and stacker-crane-based warehouses (Eder et al., 2019) (Siciliano et al., 2020).

In recent years, another warehouse was investigated that contemplates the simultaneous use of shuttles and a stacker crane. This warehouse is denoted as autonomous shuttles and stacker crane (AS/SC) warehousing system. Its shuttles move orthogonally to the stacker crane's aisle and therefore can only use the Last In First Out (LIFO)

policy. In addition, so far only one stacker crane per aisle has been implemented. (Wang et al., 2020)

On the contrary, the shuttles of a DHPW can move in both directions of the plane and up to three stacker cranes per aisle have been coordinated and investigated in (Siciliano et al., 2022). To increase the throughput of a DHPW, specific order assignment strategies (Siciliano and Fottner, 2021) and specific stacker cranes' coordination policies (Siciliano et al., 2022) should be applied that take into consideration the complex nature of the connection between shuttle tier and multiple stacker cranes in a single aisle. To investigate the nature of the connection between shuttles and stacker cranes in more detail, and to find further applications for DHPWs, this article examines two additional warehouse arrangements, which we define as layout 2 and layout 3. We call the original DHPW with the channel storage above the shuttle base layout 1. In the following section, we describe the characteristics of layout 2 and layout 3 compared to layout 1.

Systems under consideration

Layout 2 and layout 3 have shuttle tiers on not only the base but also on the levels (Malik 2014), see Fig.1.

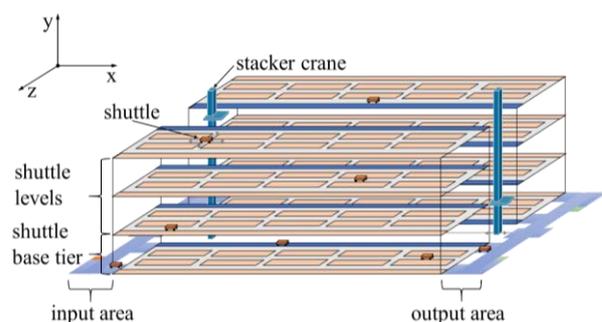


Figure 1: Structure of both Layouts 2 and 3

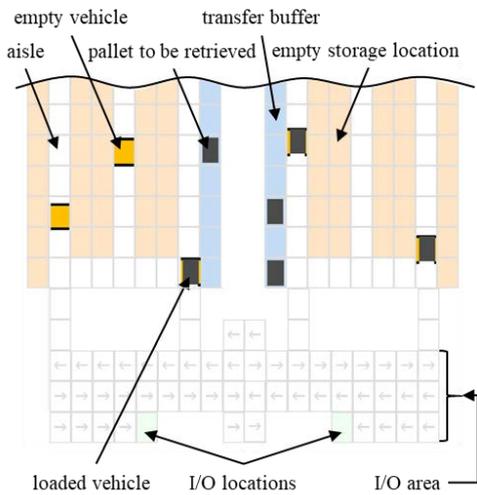


Figure 2: Screenshot of the Base Tier Model for both Layouts 2 and 3

Fork lift stacker cranes serve the transfer buffers of base tier and of the levels. Each shuttle remains in its zone i.e. left or right side of the aisle on a certain level in the warehouse. In layout 2, the shuttles cannot leave their level, while in layout 3 the shuttles can be transported by the stacker cranes between levels. The elements that make up the base tier of layout 2 and 3 are shown in Fig.2. The levels contain the same elements as the base tier, except for the fact that they lack input/output (I/O) areas. On one hand, having shuttle tiers on every level increases the investment and operational costs compared to layout 1. On the other hand, it enables better access to stored products compared to channel storage. Therefore, layout 1 can be seen as the result of the hybridization of a stacker crane-based warehouse through shuttles, while layout 2 and layout 3 are the hybridization of a shuttle-based warehouse through stacker cranes. Compared to a conventional shuttle-based warehouse with lifts, the stacker cranes' aisles in layout 2 and layout 3 offer a much more efficient means of material exchange between the base and upper levels. In fact, the transfer buffers on the base and on all levels along the whole length of the aisle provide many more exchange locations than the conventional few I/O locations of lifts. Thus, layouts 2 and 3 can achieve a higher throughput than conventional shuttle-based systems. In the following section we propose control strategies for layout 2 and layout 3.

CONCEPT DEVELOPMENT

To explain the control algorithms that were developed, we have to consider layout 2 and layout 3 separately. We implemented the algorithms in the cases of retrieval, storage and double cycles. A double cycle is the alternation of retrieval and storage orders for the shuttles. The same is true for the stacker cranes in the aisle. Therefore, retrieval and storage control strategies can be derived from the strategy for double cycles. We only discuss double cycles for the sake of brevity.

Layout 2

We first consider layout 2. The control strategy we developed for this in the case of double cycles is described in Fig. 3 for the shuttles on the base, in Fig. 4 for the shuttles on the levels, and in Fig. 5 for the stacker cranes. Abbreviations “ C_nS ” and “ C_nE ” indicate respectively start and end of connection n between shuttles and stacker cranes. The challenge compared to layout 1 is to connect and coordinate the stacker cranes with the shuttles on not only the base, but also on the different levels. For sake of completeness, we illustrate the connections between shuttles and stacker cranes for the execution of a double cycle. First, a shuttle on base executes a storage order by bringing a pallet from the I location to an available location of the transfer buffer. The shuttle then creates a storage order for the stacker crane to transport that pallet from the transfer buffer on the base to the transfer buffer of the target level, where it will be stored. The creation of such an order represents the start of one of four connection points between the control system of shuttles and that of stacker cranes. The end of connection is represented in the logic of the stacker crane by the examination of the availability status of the stacker crane. In case a stacker crane is available the storage order is executed and the pallet is delivered to the transfer buffer of the target level. At this point, the stacker crane creates a storage order for the shuttles on that level. This constitutes the start of another connection between shuttles and stacker cranes.

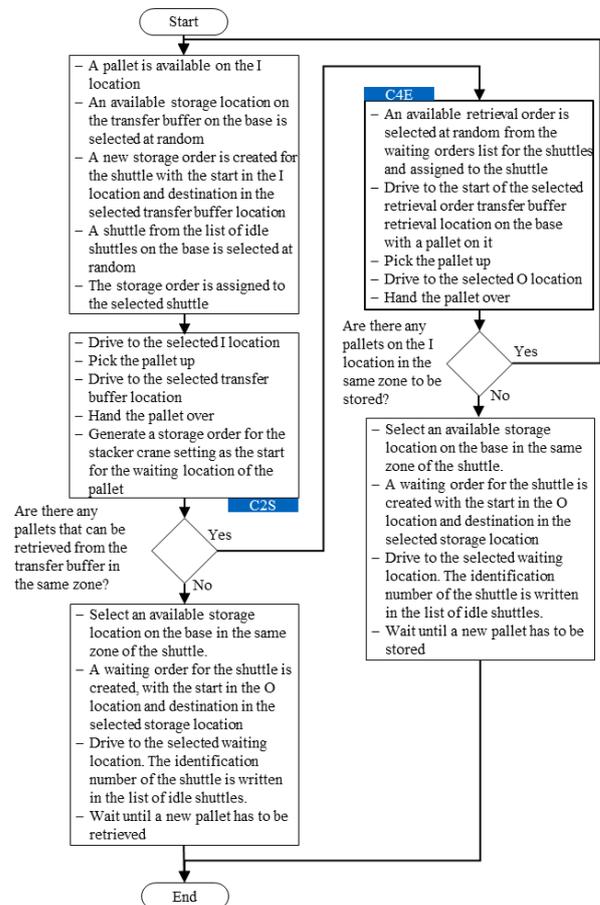


Figure 3: Control Logic – Layout 2, Double Cycles, Shuttles on Base

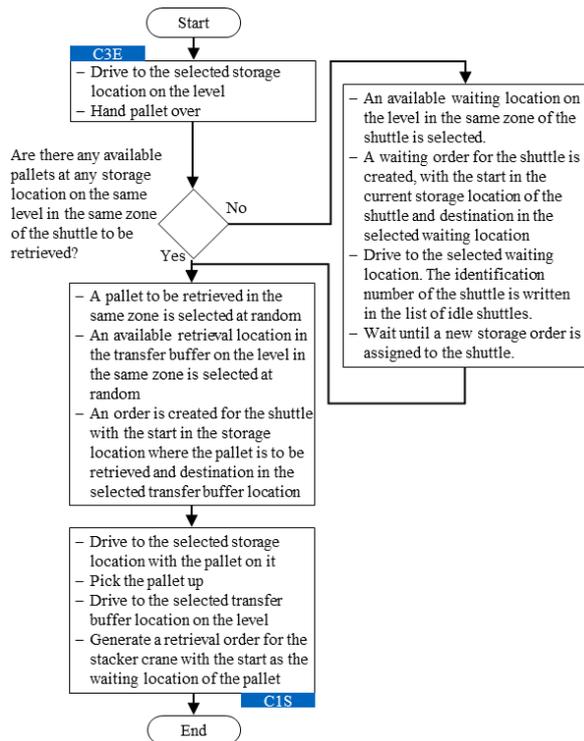


Figure 4: Control Logic – Layout 2, Double Cycles, Shuttles on Level

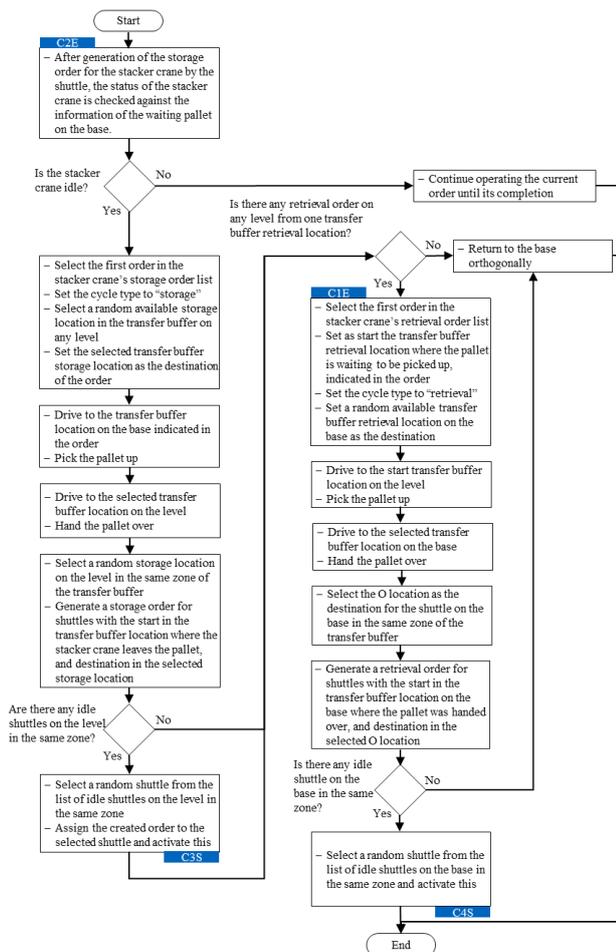


Figure 5: Control Logic – Layout 2, Double Cycles, Stacker Cranes

The end of the connection is in the logic of shuttles on level and is represented by the shuttle starting its route to pick the pallet on the transfer buffer and bring it to its final storage place. In the meantime, a shuttle on a level executes a retrieval order by moving a target pallet from its storage location to the transfer buffer. The shuttle generates then a retrieval order for the stacker crane. This generation is another connection between shuttles and stacker cranes. The end of connection is constituted by the stacker crane examining if there are retrieval orders to be executed. Next, the stacker crane performs the retrieval order by transporting the pallet to an available location of the transfer buffer on base. At this moment, another connection between stacker cranes and shuttles starts when the stacker crane generates a retrieval order for the shuttles on base. The end of connection is represented by the shuttle on base starting its route to execute the retrieval order. The shuttle transports the pallet from the transfer buffer to the O location and the double cycle is completed.

Layout 3

We now examine layout 3. The control algorithms we generated for the double cycles process in layout 3 is explained in Fig. 6 for the shuttles on the base, in Fig. 7 for the shuttles on the levels, and in Fig. 8. for the stacker cranes. The challenge, as opposed to layout 2, lies in the generation and correct assignment of transportation orders for the stacker crane to move shuttles between the different levels and the base, and of motion orders for the

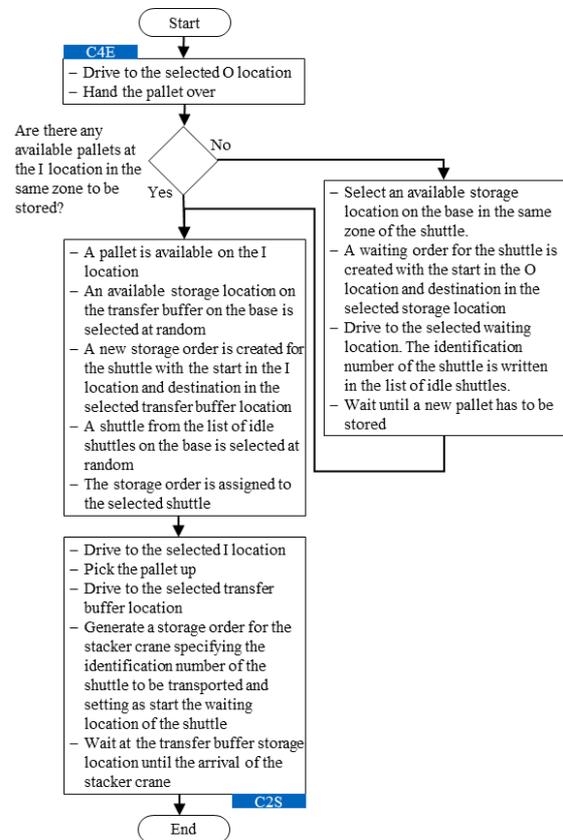


Figure 6: Control Logic – Layout 3, Double Cycles, Shuttles on Base

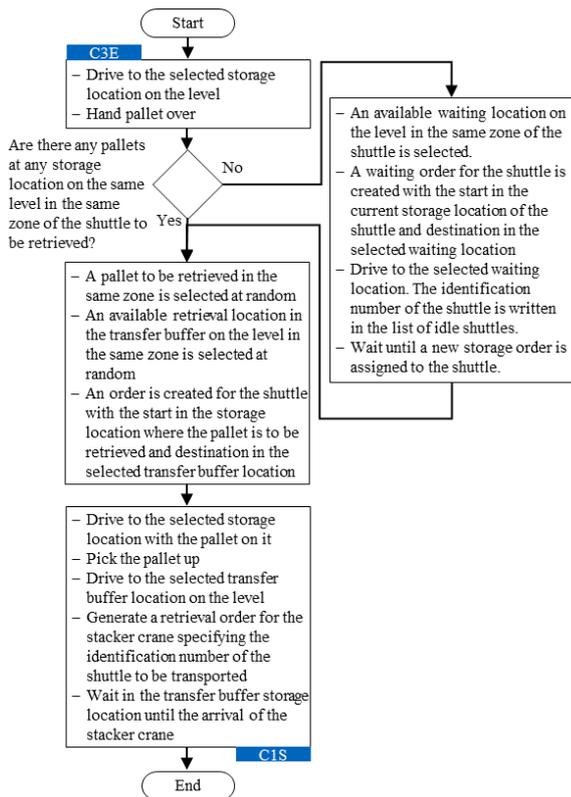


Figure 7: Control Logic – Layout 3, Double Cycles, Shuttles on Level

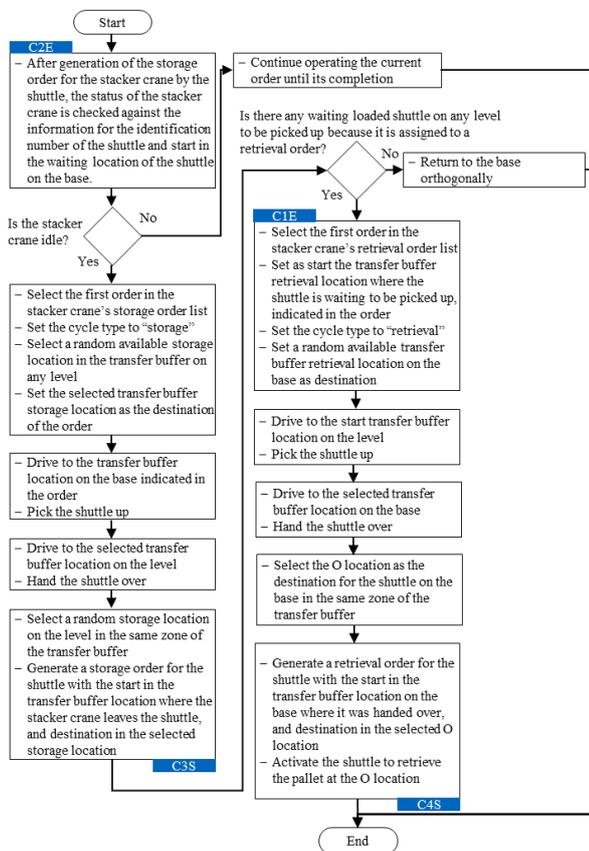


Figure 8: Control Logic – Layout 3, Double Cycles, Stacker Cranes

shuttles themselves. In fact, transportation orders should be created for not only loaded but also empty shuttles that have to be brought back from base to levels; motion orders should also be generated when an empty shuttle has to move, even if no pallet is to be picked or delivered. If the stacker cranes prove to be the bottleneck in the system, the shuttles without orders can wait directly on the transfer buffer. This saves travel time, when the stacker cranes are finally ready to exchange pallets, and energy, compared to having to drive to a waiting position in the storage locations as in layout 2.

SIMULATION STUDY

We implemented the model for layout 2 and layout 3 using the discrete event simulation environment Plant Simulation. To avoid deadlocks, the route of the shuttles on the different levels is based on the reservation of time windows, exactly like the shuttles on the base tier for layout 1 (Siciliano et al., 2020). This concept was initially introduced by (Kim and Tanchoco, 1991) and then further developed for shuttle fleets by (Lienert and Fottner, 2017a). An extensive description of the routing algorithm used for the shuttles on the levels can be found in (Lienert and Fottner, 2017b) (Lienert et al. 2020).

Parameters

The system we consider for layout 2 and layout 3 has two stacker cranes in a single aisle. We define a section as the area of the base or of a level comprised of two cross aisles. The different lengths of the aisle under consideration are two (38 m), three (54 m), four (68 m), five (83m) or ten (159 m) sections. There are three shuttle tier levels above the base. Both sides of the base have I/O area for pallets entering and leaving the warehouse. Each I/O area has two I/O locations, as shown in Fig. 2. The arrangement of cross aisles and storage aisles on the base is the same as in layout 1, see (Siciliano and Fottner, 2021), except for the I/O area. In fact, we discovered through experiments that the I/O area proposed in (Siciliano et al., 2020) creates an asymmetry in the dynamics of the shuttles for the right side of the warehouse compared to the left side for layouts 2 and 3. We therefore modified this as shown in Fig. 2 to guarantee symmetry, in other words the same performance for the right and left side of the warehouse, which resulted in an increased throughput.

The parameters used for the stacker crane in Tab. 1 and for the shuttles in Tab. 2 are provided by a manufacturer. Each experiment lasts 24 hours. We verified the model by comparing the analytically obtained travel time of individual vehicles with the simulated values (Siciliano et al., 2020). We then validated the travel time of the stacker crane and shuttles by comparing them with the values measured on the real subsystems, calculating the test positions of shuttles by the method in (Siciliano et al., 2021).

In the evaluation, we compare the throughput of layout 2 and layout 3 with following systems:

- Layout 1 with three channel storage levels above the base.

- Stacker crane-based warehouses whose throughput values are provided by a manufacturer.
- Shuttle-based warehouses, which we simulated in Plant Simulation. To make this comparable with DHPWs, we used the same shuttle tiers as for layout 2 and layout 3. The system has a total of four lifts i.e. two for each side of the warehouse, these being located at one third and two thirds of the length of the aisle.

The abbreviations used for the different systems examined from Fig. 9 to Fig. 14 are explained in the list of abbreviations at the end of this article.

Table 1: Stacker Crane Parameters

Parameter	Value
Speed (loaded)	0.6 m/s
Speed (empty)	1.0 m/s
Acceleration (loaded)	0.3 m/s^2
Acceleration (empty)	0.6 m/s^2
Turning time	6.6 s
Handover time	10.0 s

Table 2: Shuttle Parameters

Parameter	Value
Travel speed x	4.0 m/s
Travel acceleration x	0.5 m/s^2
Lifting speed y	1.0 m/s
Lifting acceleration y	1.0 m/s^2
Time of pallet handover	6.0 s
Time for positioning before channel	1.0 s

Evaluation

We first studied the throughput of layout 2 by varying the length of the aisle. For both of the processes of retrieval (Fig. 9) and of double cycles (Fig. 10), reducing the length of the aisle reduces the travel distance for the shuttles, resulting in an increase in throughput. However, this increase is not particularly high, so we can conclude that the length of the warehouse does not have a great influence on the throughput. It is significant for the scalability of the system, that up to a total of 64 shuttles, the shuttles remain the bottleneck in terms of the performance of the system for retrieval and double cycles. This means that simply increasing the number of shuttles would result in a further increase in throughput.

We now consider the behaviour of layout 3 as the length of the aisle changes. For the process of retrieval in Fig. 11, as for that of double cycles in Fig. 12, the length of the aisle has less influence on the throughput than in layout 2. Moreover, layout 3 allows a higher throughput than in

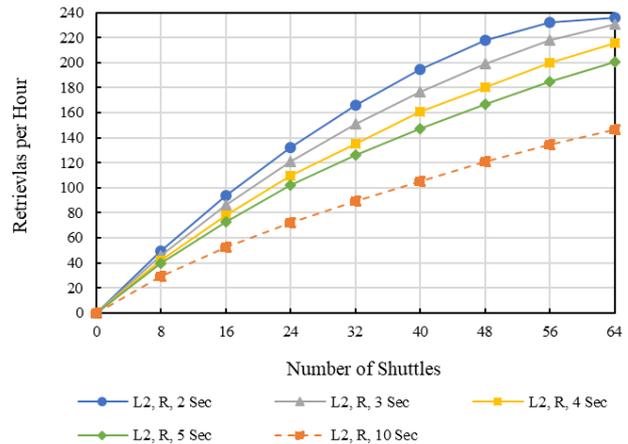


Figure 9: Retrieval Performance of Layout 2 Varying the Length of the Aisle from 2 Sections to 10 Sections

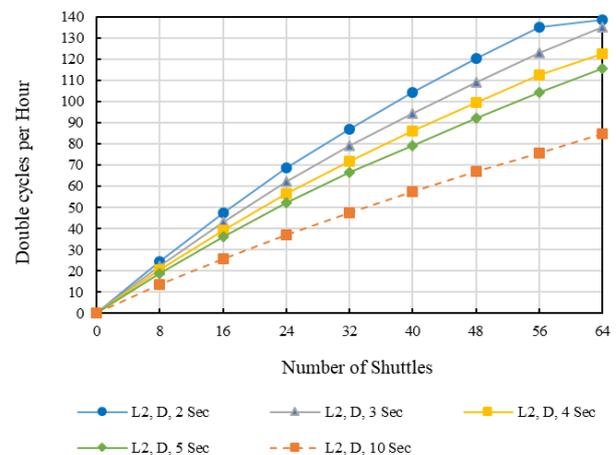


Figure 10: Double Cycles Performance of Layout 2 Varying the Length of the Aisle from 2 Sections to 10 Sections

layout 2 with a smaller number of shuttles. However, in the case of retrieval, layout 3 is limited by the bottleneck due to the stacker cranes, indicated by the plateaux in the curves, with a smaller number of shuttles than layout 2. Once the bottleneck of the stacker cranes is reached in layout 3, additional shuttles do not increase the throughput. Therefore, layout 3 is less scalable than layout 2. In Fig. 11 and Fig. 12 the results for 48 or more shuttles by two sections of layout 3 are not reported, because we do not recommend to use such a high number of shuttles in this case. The reason is that, when shuttles are able to change their levels, 48 or more shuttles are too many for the short layout of two sections and this causes congestions of shuttles near the transfer buffer of the base. As a consequence, throughput is reduced. This is a further demonstration of the lower scalability of layout 3 compared to layout 2. Not only do we compare layout 2 and layout 3 with each other, but also with other warehouses, as in Fig. 13 for retrieval and in Fig. 14 for double cycles. Layout 3 with four shuttles per level has a

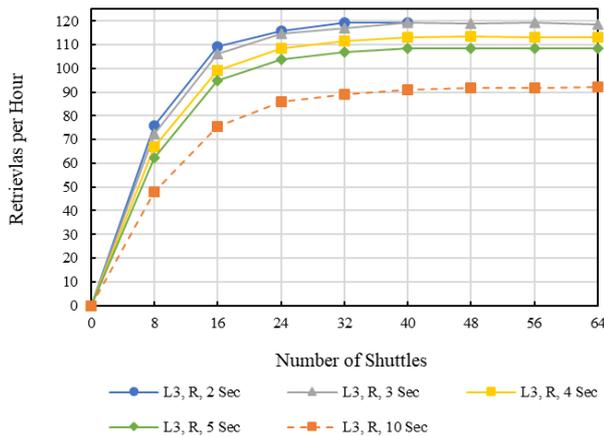


Figure 11: Retrieval Performance of Layout 3 Varying the Length of the Aisle from 2 Sections to 10 Sections

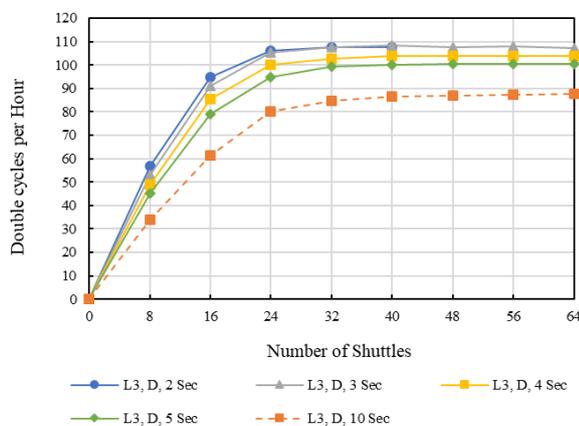


Figure 12: Double Cycles Performance of Layout 3 Varying the Length of the Aisle from 2 Sections to 10 Sections

throughput that is already higher than those of the other systems, in the case of both retrieval and double cycles.

By comparison, layout 2 needs up to 6 shuttles per level to provide a throughput that is not only higher than that of conventional stacker cranes but also of that of the shuttle-based warehouse with four lifts, which has comparable costs. Layout 1 achieves a higher throughput than that of conventional stacker crane-based warehouses, but one that is inferior to that of layout 2 and layout 3. The reason using stacker cranes, as in layouts 2 and 3, improves performance compared to using lifts is that the interface between lifts and shuttle tiers is made up of a reduced number of locations on the transfer buffer. Therefore, shuttles wait longer for a location or a pallet to become available than in the case of stacker cranes, which have locations on the transfer buffer all along the aisle.

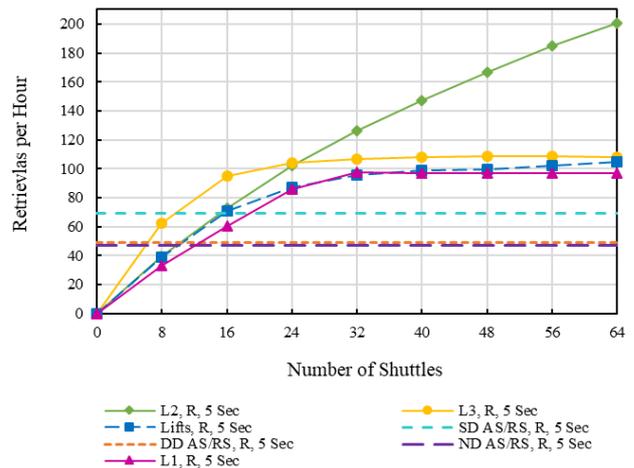


Figure 13: Comparison of Retrieval Performance between Different Warehouse Systems

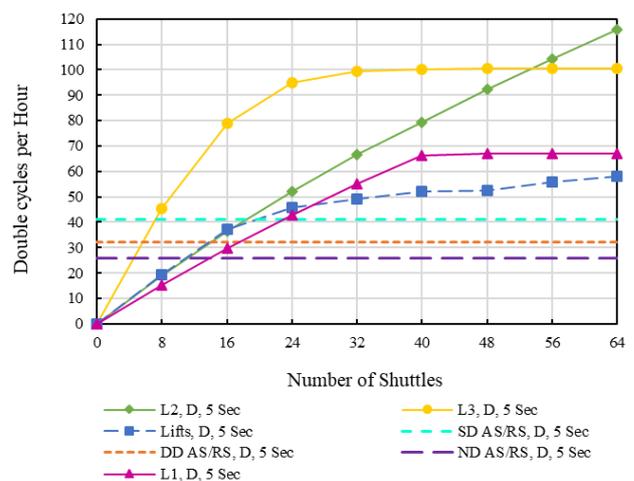


Figure 14: Comparison of Double Cycles Performance between Different Warehouse Systems

CONCLUSION AND OUTLOOK

In this article, we described the design and advantages of two warehouses, classifiable as DHPWs, which we call layout 2 and layout 3 respectively. We then proposed control strategies for each of these. Through a discrete event simulation study, we demonstrated that the length of the stacker cranes' aisle has no great influence on the throughput for either of them. With a small number of shuttles, layout 3 should be given preference over layout 2 because its poorer scalability is not yet dominant: moving shuttles to the levels where they are needed more urgently then overcompensates the additional orders for the stacker cranes. When using many shuttles, these additional orders lead to an earlier bottleneck, so that layout 2 is then preferable. All in all, the results of this paper are decision support in warehouse management for pallets insofar as they illustrate the performance benefits of substituting layout 1 to multi-depth stacker crane-based warehouses applications and of replacing the connection

to lifts in shuttle-based warehouses with a connection between shuttles and stacker cranes such as layouts 2 and 3. For future research, different coordination algorithms between shuttles and stacker cranes have to be investigated to further improve the throughput without having to increase the number of shuttles.

ACKNOWLEDGEMENTS

We are very grateful to Thomas Klopfenstein from the firm of Gebhardt Fördertechnik GmbH for providing us with the throughput values for conventional stacker crane-based warehouses.

Moreover, we would like to thank Joerg Eder, also from the firm of Gebhardt Fördertechnik GmbH, for the fruitful collaboration.

REFERENCES

- Eder, J., Klopfenstein, T., & Gebhardt, M., 2019. *Patent: Lagersystem zur Speicherung und Abgabe von Ladungsträgern*. DE102019211804, German Patent and Trade Mark Office (DPMA).
- Kim, C. W., & Tanchoco, J. M. A., 1991. Conflict-free shortest-time bi-directional AGV routing. *International Journal of Production Research* 29 (12): 2377-2391.
- Lienert, T., & Fottner, J., 2017. No more deadlocks—applying the time window routing method to shuttle systems. *Proceedings of the 31st European Conference on Modelling and Simulation (ECMS)*, 169–175.
- Lienert, T., & Fottner, J., 2017. Development of a generic simulation method for the time window routing of automated guided vehicles. *Logistics Journal: Proceedings*, Vol. 2017.
- Lienert, T., Wenzler, F., & Fottner, J., 2020. Simulation-based evaluation of reservation mechanisms for the time window routing method. *Proceedings of the 33rd European Conference on Modelling and Simulation (ECMS)*.
- Malik, O., 2014. *Patent Application Publication: Automated warehousing systems and method*. US20140086714A1, US Patent and Trademark Office (USPTO).
- Siciliano, G., Lienert, T., Fottner, J., 2020. Design, Simulation and Performance of a Highly-Dynamic, Hybrid Pallet Storage and Retrieval System. *Proceedings of the 19th International Conference on Modeling and Applied Simulation (MAS)*.
- Siciliano, G. & Fottner, J., 2021. Concept development and evaluation of order assignment strategies in a highly dynamic, hybrid pallet storage and retrieval system. *Proceedings of the 11th International Conference on Simulation and Modeling Methodologies (SIMULTECH)*, ISBN 978-989-758-528-9, ISSN 2184-2841, pp. 360-368.
- Siciliano, G., Durek-Linn, A., Fottner, J., 2022. Development and Evaluation of Configurations and Control Strategies to Coordinate Several Stacker Cranes on a Single Aisle for a New Dynamic Hybrid Pallet Warehouse. In: Shi X., Bohács G., Ma Y., Gong D., Shang X. (eds) *LISS 2021. Lecture Notes in Operations Research*. Springer, Singapore. https://doi.org/10.1007/978-981-16-8656-6_54
- Siciliano, G., Schuster, C. U., Fottner, J., 2021. Analytical method to determine the test positions for validation of a two-dimensional shuttle system model. *Proceedings of the 20th International Conference on Modeling & Applied Simulation (MAS)* 2021), pp. 21-28. <https://doi.org/10.46354/i3m.2021.mas.003>
- Wang, Y., Man, R., Zhao, X., Liu, H., 2020. Modeling of parallel movement for deep-lane unit load autonomous shuttle and stacker crane warehousing systems. *Processes* 8(1)

LIST OF ABBREVIATIONS

- L1 = Layout 1; L2 = Layout 2; L3 = Layout 3
- R = Retrieval process; D = Double cycles process
- Sec = Sections
- SD AS/RS = Single-deep storage stacker crane with telescopic forks
- DD AS/RS = Double-deep storage stacker crane with telescopic forks with relocations
- ND AS/RS = Nine-deep storage stacker crane with satellite without relocations

AUTHOR BIOGRAPHIES

GIULIA SICILIANO is a research associate at the Chair of Materials Handling, Material Flow and Logistics in the School of Engineering and Design at Technische Universität München. She holds a M.Sc. in Mechanical Engineering from Università degli Studi di Roma “Tor Vergata”. Her research interests lie in the design, control and simulation of automated warehouses, and in the development of artificial intelligence models for optimization of logistical processes.

YUE YU is a former master student of the Chair of Materials Handling, Material Flow and Logistics in the School of Engineering and Design at Technische Universität München and now process simulation engineer at Dräxlmaier Group. She holds a M.Sc. in Development, Production and Management. Her research interests include the design and simulation of automated warehouses.

JOHANNES FOTTNER is Professor and Head at the Chair of Materials Handling, Material Flow, Logistics in the School of Engineering and Design at Technische Universität München.

EIN FUNKTIONALER VERGLEICH DER SAP ANALYTICS CLOUD UND MICROSOFT POWER BI ZUR VERWENDUNG IM BEREICH PEOPLE ANALYTICS BEI VITESCO TECHNOLOGIES

Fabian Engl
Ostbayerische Technische
Hochschule Regensburg,
Vitesco Technologies
People Analytics and Technology
Siemensstraße 10-12,
93055 Regensburg
Email: fabian.engl@vitesco.com

Philipp Trubjansky
Vitesco Technologies
People Analytics and Technology
Siemensstraße 10-12,
93055 Regensburg
Email: philipp.trubjansky@vitesco.com

Professor Dr. Frank Herrmann
Ostbayerische Technische
Hochschule Regensburg
Labor Wirtschaftsinformatik, SAP
und Produktionslogistik
Galgenbergstraße 32,
93053 Regensburg
Email: frank.herrmann@oth-regensburg.de

Abstract—Many companies, including Vitesco Technologies' People Analytics department, are using Business Intelligence (BI) solutions to distinguish upcoming trends and support strategic decision making. Utilizing BI-Software has expanded beyond finance and controlling departments. Cross-department utilization and acceptance along with trends like Cloud Computing and Big Data becoming the new norm, have shaken the BI software landscape. Solutions that were implemented in the past cannot always cope with these new developments. Therefore, especially specific BI application areas like People Analytics need regular software re-evaluations.

This article briefly highlights the characteristics of Cloud Business Intelligence, identifies new challenges for People Analytics and proposes a BI software comparison model based on technical capabilities to be used within People Analytics at Vitesco Technologies. Following, the current software – Power BI – will be compared to the cloud-based SAP Analytics Cloud. Conclusively a recommendation for a future People Analytics BI software setup will be given.

I. EINFÜHRUNG

Business Intelligence gewinnt in den letzten Jahren immer mehr an Bedeutung. Laut einer Umfrage des Contact-Center-Networks, gaben 2020 knapp 44 Prozent aller Befragten Unternehmen an, sie planen dieses Jahr bewusst in Business Intelligence investieren zu wollen. Dies entspricht einem deutlichen Zuwachs im Vergleich zu den Vorjahren (Contact-Center-Network 2020). Eine weitere Studie von Statista belegt diesen Trend. Ihr zufolge wuchs der globale Markt für Business-Intelligence-Software zwischen 2020 und 2021 um 4,5 Prozent auf einen Gesamtwert von 23 Milliarden Euro. Bis 2026 prognostiziert die Studie einen weiteren Zuwachs von über 43 Prozent. Microsofts Power BI und SAPs Analytics Cloud teilen sich nach aktuellem Stand mit acht Prozent Marktanteil den zweiten Platz im weltweiten Business-Intelligence-Softwaremarkt. Den ersten Platz belegt das Unternehmen SAS (Statista Technology Market Outlook 2021).

Vitesco Technologies, insbesondere People Analytics, verwendet aktuell die Business-Intelligence-Software Power BI. Mithilfe dieser werden klassische Key-Performance-Indicators (KPIs) des Personalwesens wiedergegeben, um im Human-Resources-Umfeld (HR) Hilfe bei Personalentscheidungen zu leisten.

Dieser Artikel vergleicht Power BI und die mögliche Alternativsoftware SAP Analytics Cloud auf funktionaler Basis. Nach einer genaueren Beschreibung der Problemstellung, werden zunächst für den Vergleich benötigte grundlegende Definitionen gegeben sowie Eigenschaften und Besonderheiten einer Cloud-Business-Intelligence-Software erarbeitet. Danach folgt eine Identifikation von aktuellen Herausforderungen im Bereich People Analytics und die Auswahl eines geeigneten Vergleichsmodells. Abschließend werden ein Fazit und eine Handlungsempfehlung ausgesprochen.

II. PROBLEMBESCHREIBUNG

People Analytics liefert als *Single Source of Truth* eine zentrale Anlaufstelle für Fragen rund um das eigene Personal und dessen Entwicklung. Hauptinteressenten sind dabei neben dem oberen Management vor allem HR-Mitarbeiter mit Länder- sowie Führungskräfte mit weitreichender Personalverantwortung. Durch die zahlreichen Stakeholder und ihren verschiedenen fachspezifischen Erwartungen müssen zentrale Berichte vielen Anforderungen gerecht werden. Das Anpassen der Berichte und deren Datenbasis nimmt dabei viel Zeit in Anspruch. Die verwendete Software Power BI trägt durch funktionale und leistungsbedingte Einschränkungen stark hierzu bei.

Somit stellt sich die Frage, ob Power BI in diesem Zusammenhang trotzdem die geeignetste Business-Intelligence-Software für den Einsatz im Bereich People Analytics bei Vitesco Technologies darstellt. Durch eine zeitgleiche interne Evaluation der SAP Analytics Cloud (SAC) zur Verwendung in anderen Unternehmensbereichen erscheint es aus wirtschaftlicher Sicht sinnvoll, diese als Alternative für den Vergleich heranzuziehen. Mögliche andere Anbieter wurden im Vorfeld zwar überprüft, allerdings aufgrund interner Entwicklungen und Entscheidungen vom weiteren Vergleich ausgeschlossen.

Ziel ist eine Überprüfung der SAC auf die speziellen Anforderungen des internen People-Analytics-Teams und die Identifikation von Herausforderungen eines möglichen Umstiegs. Letzteres erfolgt besonders unter dem Aspekt einer Verlagerung in die Cloud.

III. BEGRIFFSDEFINITIONEN

A. *Cloud Business Intelligence*

Cloud Business Intelligence beschreibt eine Kombination aus Cloud Computing und Business Intelligence. Um das Konzept zu erklären, ist zuerst eine Ausführung der Begriffe Business Intelligence und Cloud Computing nötig.

1) *Business Intelligence:*

Der Begriff Business Intelligence wurde erstmalig im Jahr 1865 von Richard Millar Devens in seinem Buch *Cyclopaedia of Commercial and Business Anecdotes* erwähnt (Devens 1865). Erst 1996 definierte die Gartner Gruppe, ein bis heute führendes Unternehmen im Bereich IT-Trend- und Marktforschung (Gartner Inc. 2021), Business Intelligence erstmalig im Kontext der heute etablierten Informations- und Kommunikationssysteme (Zeng u. a. 2006). Heute versteht man unter dem Begriff Business-Intelligence Systeme und Anwendungen, die das obere Management unmittelbar bei der Entscheidungsfindung unterstützen (Kemper, Baars und Mehanna 2010; Gluchowski, Gebriel und Dittmar 2008). Business Intelligence beschreibt also die Gesamtheit der „Verfahren und Technologien [...], die interessante Muster in umfangreichen Datenbeständen aufdecken und Prognosen über zukünftige Ereignisse und Gegebenheiten anstellen können“ (Gluchowski 2016). Durch dieses technische Vorgehen unterstützt Business Intelligence maßgeblich bei der strategischen Unternehmensführung sowie der Entscheidungsfindung und gibt einen Überblick über Leistungen und Effektivität von Unternehmensbereichen sowie deren Geschäftsprozesse (Hostmann, Rayner und Friedmann 2006).

2) *Cloud Computing:*

Zu Cloud Computing existieren in der Fachliteratur verschiedene Definitionen, immer wieder wird jedoch auf die Definition des National Institute of Standards and Technology verwiesen (Gurjar und Rathore 2013; Satyanarayana 2012): „[C]loud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction“ (National Institute of Standards and Technology’s 2011). Cloud Computing lässt sich in diesem Zusammenhang in folgende drei Servicekategorien einteilen (Satyanarayana 2012):

- **Software-as-a-Service (SaaS):** Die Software ist als Service verfügbar, die technische Umsetzung obliegt dem Serviceprovider.
- **Platform-as-a-Service (PaaS):** Der Endanwender erhält eine Plattform für seine Anwendungen.
- **Infrastructure-as-a-Service (IaaS):** Der Serviceprovider stellt lediglich die technischen Komponenten wie Rechenleistung, Netzwerkbandbreite und Speicherkapazität bereit (Olszak 2014).

Diese Servicekategorien können als Private, Community, Public und Hybrid Clouds angeboten werden. Eine Private Cloud ist im Besitz eines einzelnen Unternehmens,

während auf Community oder Public Clouds andere Unternehmen bzw. die Öffentlichkeit Zugriff haben. Eine Hybrid Cloud beschreibt die Kombination aus zwei oder mehreren verschiedenen Cloud-Modellen. Der Endnutzer muss sich bei keinem Modell Spezialwissen über die geleastete Software, Plattform oder Infrastruktur aneignen, da diese vom Anbieter bereitgestellt werden. Der Steuer- und Kontrollaufwand auf Endnutzerseite entfällt vollständig (Savu 2011; Mircea, Ghilic-Micu und Stoica 2011).

3) *Cloud Business Intelligence:*

Man spricht also von Cloud Business Intelligence, wenn eine oder mehrere Komponenten des Gesamtsystems als eigenständige, unabhängige Services über die Cloud bereitgestellt werden. Dazu zählt bereits die Datenspeicherung in einer Cloud. Data Warehouses, Datenverarbeitungsprozesse oder die Business-Intelligence-Anwendung bilden weitere mögliche Komponenten. Cloud- und klassische On-Premise-Systeme können beliebig kombiniert werden. Das ermöglicht kritische Abläufe und Daten weiterhin auf firmeninternen Servern zu halten, während das Gesamtsystem an anderen Stellen die Vorteile von *Cloud Computing* nutzt (Muntean 2015).

B. *People Analytics*

People Analytics und Business Intelligence stehen sich als Fachdisziplinen sehr nahe. People Analytics beschreibt dabei speziell die „systematische Analyse von Daten aus dem Personalwesen in Verbindung mit Daten aus anderen Unternehmensbereichen mit dem Ziel, Faktoren der Zusammenarbeit von Mitarbeitern und der Wettbewerbsfähigkeit von Unternehmen besser zu verstehen und gezielt zu fördern“ (C. U. Reindl 2016). Die rein datengestützten Erkenntnisse werden insbesondere durch Theorien aus der Sozial-, Emotions- und Motivationspsychologie ergänzt (C. U. Reindl 2016). Zu den Aufgabenbereichen von People Analytics gehören sowohl standardisierte Reports zu grundlegenden Mitarbeiter- und Unternehmenskennzahlen als auch analytisch fundierte Prognosen über mögliche zukünftige Entwicklungen innerhalb des Unternehmens (Peeters, Paauwe und Van De Voorde 2020; Isson und Harriott 2016). Je nach Grad der Informationsnutzung können verschiedene Ausprägungsstufen erreicht werden, diese beginnen beim reaktiven Reporting und reichen bis zur vollständigen Automatisierung von Entscheidungen (Holthaus, Park und Stock-Homburg 2015).

IV. BESONDERHEITEN VON CLOUD-BUSINESS-INTELLIGENCE-ANWENDUNGEN IM BEREICH PEOPLE ANALYTICS

A. *Eigenschaften einer Cloud-Business-Intelligence-Software*

Durch die zunehmend komplexere Unternehmenswelt sehen sich Unternehmen gezwungen, kontinuierlich innovative Dienstleistungen anzubieten, um wettbewerbsfähig zu bleiben. Für Business Intelligence äußern sich diese Entwicklungen primär in der Anforderung, mehr quantitative Daten zu verarbeiten. Die Fähigkeit, diese Daten zu strukturieren und basierend darauf zu handeln, hat direkten Einfluss auf die Wettbewerbsfähigkeit eines Unternehmens.

Konventionelle Business-Intelligence-Lösungen sind meist nicht auf eine derartige Datenflut ausgelegt. Oft scheitern kompliziertere analytische Anforderungen an den starren Systemstrukturen (Gurjar und Rathore 2013). Des Weiteren erfordert die Einführung einer traditionellen Business-Intelligence-Software hohe finanzielle Investitionen und große personelle Kapazitäten. Serviceorientierte Lösungen sind hier deutlich flexibler und agiler (Mircea, Ghilic-Micu und Stoica 2011). Flexibilität wird bei Cloud Business Intelligence unter anderem durch die Bereitstellung der Endanwendung in Form einer SaaS-Lösung und einer skalierbaren IT-Architektur seitens des Softwareanbieters erreicht (Liu u. a. 2010).

Cloud-Business-Intelligence-Lösungen sind dabei nicht zwingend für alle Unternehmen geeignet. Die erfolgreiche Einführung hängt von vielen verschiedenen Faktoren ab (Mircea, Ghilic-Micu und Stoica 2011). Trotz vieler Herausforderungen verwendete laut Statista 2021 bereits jedes achte Unternehmen in Deutschland eine Form von Cloud Business Intelligence, doppelt so viele wie im Vorjahr. Vor allem international tätige Unternehmen setzen vermehrt auf Cloud-Lösungen (Capgemini 2021).

B. Einfluss von Cloud-Business-Intelligence-Software auf den Bereich People Analytics

People Analytics hat ebenfalls mit einer täglich steigenden Menge an Daten und deren Vielfalt zu kämpfen (Sharda, Delen und Turban 2015; Gurjar und Rathore 2013). Wie erfolgreich die eigenen Personalstrategien und -maßnahmen sind, hängt immer mehr von der Fähigkeit, konzerneigene Datensätze sinnvoll mit diesen externen Datenquellen zu erweitern, ab (Isson und Harriott 2016). Die Kombination aus internen und externen Daten erlaubt es, aufschlussreiche und relevante Erkenntnisse über das eigene Personal zu erlangen (Kavis 2014). Ein Großteil dieser im Internet verfügbaren Daten ist unstrukturiert und kann von einer konventionellen Business-Intelligence-Software nicht ohne weitere Zwischenverarbeitungsschritte dargestellt werden (Gurjar und Rathore 2013).

Um diese relevanten Datenquellen vollumfänglich auszuwerten und anzuzeigen, benötigen Systeme in der Theorie unendlich viele technische Ressourcen. Da keine Organisation über diese Mengen an Rechenleistung verfügt, kommen vermehrt sogenannte Queries zum Einsatz. Sie rufen eine speziell definierte, überschaubare Menge an Datensätzen im Web ab und gruppieren diese. Eine konventionelle On-Premise-Business-Intelligence-Software scheitert oft an den limitierten Ressourcen (Gurjar und Rathore 2013).

C. Einbindung von Business-Intelligence-Software in die HR-Systemlandschaft

Im Personalwesen existieren sogenannte Human-Relation-Information-Systems (HRIS). Diese wurden ursprünglich eingeführt, um HR-Prozesse einfacher und agiler zu gestalten. Außerdem erlauben sie Angestellten einen leichteren Zugang zu Personaldaten. Viele der großen HRIS-Anbieter liefern ihre Produkte mittlerweile mit integrierten Analytics-Modulen aus. Diesen spezifischen Modu-

len mangelt es allerdings oft an ausreichender Leistung für die komplexen analytischen Prognosen von People Analytics. Resultierend daraus scheitern viele Unternehmen an der Umsetzung prädiktiver analytischer Verfahren und geben lediglich historische Kennzahlen wieder (Angrave u. a. 2016).

Trotz der mangelnden analytischen Auswertungsmöglichkeiten der mitgelieferten Module bieten HRIS einen entscheidenden Vorteil für eine funktionsumfangreichere Business-Intelligence-Software, sie dienen als globale HR-Datenbank. Personaldaten sind an einem zentralen Ort abgespeichert und die hohe Datenverfügbarkeit ermöglicht umfangreiche Auswertungen (Loscher 2021). Dafür müssen die Daten allerdings auch zentral abrufbar sein. Sind sie das nicht, besteht die Gefahr, dass Abteilungen für ihre eigenen Zwecke Datenbanken aufsetzen. Darunter leidet die Qualität sowie Korrektheit der Daten (Pentzek und Espig 2016).

Aus Datenschutzgründen hat People Analytics oft selbst mit ähnlichen Problemen zu kämpfen. Der direkte Zugriff auf Personaldaten ist aufgrund verschiedener Gesetze nur erschwert möglich und erfordert in vielen Fällen Anonymisierungsmaßnahmen vor der Weiterverarbeitung der Daten. Dennoch sind Auswertungen unter Berücksichtigung der rechtlichen Regularien möglich (Wille 2016). Der Fokus dieses Vergleichs liegt auf dem rein funktionalen Umfang der Business-Intelligence-Software. Vollständigkeitshalber werden rechtliche Vorgaben wie der Datenschutz kurz aufgeführt, aber in der späteren Auswahl nicht berücksichtigt.

D. Rechtliche Vorgaben

Regulatorische Vorgaben sorgen besonders bei den Datenverarbeitungsprozessen für zusätzliche technische Anforderungen. Im Personalwesen, das stark durch personenbezogene Daten geprägt ist, kann es leicht zur Verletzung der Privatsphäre kommen. Unternehmen, damit auch People Analytics, müssen sich hier an eine Reihe geltender Gesetze halten und innerhalb dieser juristischen Grenzen agieren.

Die geltenden Gesetze sichern Mitarbeitern Transparenz über die Verarbeitung ihrer personenbezogenen Daten zu. People Analytics trägt damit die Aufgabe, „[...] diese Fakten auf einfache und wirksame Weise den interessierten Parteien zu vermitteln“ (Loi 2021). Um das zu garantieren, wird ein genauer Überblick über die verarbeiteten Daten sowie Zwischenspeicherstände nach den einzelnen Verarbeitungsschritten benötigt. Das ist besonders dann relevant, wenn entscheidungstragende Technologien wie Künstliche Intelligenz zum Einsatz kommen (Loi 2021). Für Unternehmen, die in vollem Umfang rechtskonform und offen handeln, ist es essenziell, diese Prozesse zu beschreiben und zu dokumentieren. Eine klare Kommunikation der Verarbeitungsprozesse und verwendeten Daten trägt zusätzlich zur Akzeptanz von People Analytics bei (Loi 2021).

E. Datenqualität und -struktur in HRIS

Datenqualität spielt in allen Anwendungsbereichen von Business Intelligence eine große Rolle. Im Personalwesen trägt diese dennoch maßgeblich zum Erfolg und der unternehmensweiten Akzeptanz von People Analytics bei. Viele Arbeitsschritte, wie etwa das Einpflegen von neuen Mitarbeiterdaten, erfolgen noch manuell (Loscher 2021). Die Verantwortung einer klaren und sauberen Grundstruktur obliegt somit nicht einer einzelnen Person oder Abteilung. Hohe Datenqualität kann nur erreicht werden, wenn sich alle betroffenen Mitarbeiter dieser Verantwortung bewusst sind und aktiv zu deren Aufrechterhaltung beitragen. Datenqualität sowie redundante und manuelle Datenpflege zählen zu den größten Herausforderungen des Datenqualitätsmanagements und somit auch People Analytics (Otto und Österle 2016).

Ein Mangel an Datenqualität sorgt für hohen Mehraufwand und vermehrte Fehler im Laufe der weiteren Datenverarbeitung (Pentzek und Espig 2016). Eine lückenhafte Datenbasis liefert nicht die von People Analytics gewünschten entscheidungsunterstützenden Ergebnisse. Fehlerhafte Datensätze können zwar (eingeschränkt) für die Berichterstattung herangezogen werden, sollten aber für weitere statistische Auswertungen ignoriert werden. Eine inkorrekte Eingabe führt zwangsweise zu einem schlechten Ergebnis und so im schlimmsten Fall zu falschen Entscheidungen, folglich einer grundlegenden Zweckverfehlung von People Analytics (Loscher 2021). Als Folge der ungenügenden und mangelhaften Daten können viele strategisch wertvolle Fragen nicht beantwortet werden (Angrave u. a. 2016).

F. Analytische Anforderungen im Bereich People Analytics

Analytische Auswertungen im HR-Bereich gelten als sehr komplex. In der spezifischen Fachliteratur werden hierfür selten genaue Gründe angeführt. Die großen Datensätze, wie sie im Personalmanagement oft zu finden sind, erschweren allerdings Auswertungen, denn sie benötigen mehr Rechenleistung (Gani u. a. 2016). Existierende HRIS sind zudem auf die Anforderung des Personalwesens abgestimmt und „[d]ie derzeitige IT-Infrastruktur ist in den meisten Personalabteilungen nicht für analytische Auswertungen verwendbar“ (Wille 2016). Durch den starken Fokus auf operativer Berichterstattung reicht die Datenbasis in HRIS selten aus, um strategische Fragen zu beantworten. Diese Art der Berichterstattung ist zwar von großer Bedeutung, gewinnt aber erst durch analytische Verfahren strategische Relevanz. Für die Verwendung im Bereich People Analytics reichen diese Systeme nicht aus und müssen zwingend durch Business-Intelligence-Module und -Software erweitert werden (Angrave u. a. 2016).

Viele Auswertungen übersteigen zudem klassische Daten- und Informationsanalysen, denn ermittelte Datenkorrelationen werden zusätzlich unter dem Einfluss komplexer soziologischer Modelle betrachtet. People Analytics benötigt deswegen ein wesentlich tieferes statistisches Methodenwissen, als es in anderen Business-Intelligence-

Einsatzgebieten erforderlich ist (C. Reindl und Krügl 2017).

G. Key-Performance-Indicators im Personalwesen

Als Folge des heutigen Konkurrenzdrucks sind Unternehmen dazu gezwungen, ihre wirtschaftliche Leistung zu messen und mit Wettbewerbern zu vergleichen. KPIs spiegeln diese Leistung in komprimierter und datengestützter Art wieder. Sie sind Teil der Berichterstattung und ein wichtiges Kontrollinstrument der Unternehmensführung (Iveta 2012; Filipowicz-Florczyk 2014).

Die mangelnde Datenqualität und zugleich großen Datensätze stellen People Analytics auch hier vor große Herausforderungen. Die Berechnungen von relevanten KPIs, besonders in Bezug auf zukünftige Entwicklungen, gestalten sich wegen dem Mangel an nötigen Daten oft schwer. Einige der von People Analytics untersuchten Zusammenhänge können zudem prinzipiell nur schwer berechnet werden, etwa wie sich bestimmte Schulungen auf die Produktivität der Mitarbeiter oder das Betriebsklima auswirken (C. Reindl und Krügl 2017).

Für das Personalwesen sind sie zudem nur dann förderlich, wenn eine fortlaufende konsistente Darstellung möglich ist und sie zugleich einen aussagekräftigen Einblick in das gesamte Unternehmen bieten (Iveta 2012). Eckerson schreibt im Business Intelligence Journal, KPIs sollten mit Bedacht und in Maßen abgebildet werden – „The fewer KPIs, the better“ (Eckerson 2009). Die korrekte Auswahl dieser Kennzahlen ist in Anbetracht der riesigen Datenmengen, deren Relationen und der Anzahl an Quellsystemen sehr komplex. Um sie zu erleichtern, wurden in der ISO 30414 standardisierte KPIs für das *Human Capital Reporting* festgelegt und in verschiedene Kategorien eingeteilt. Dazu gehören unter anderem die Bereiche *Diversity*, *Productivity*, *Recruitment*, *Succession planning* und *Workforce availability*. Insgesamt sind über 40 KPIs (Stand 2018) samt Beschreibung und Berechnungsgrundlage aufgeführt. Zu den wichtigsten, auch oft in der Literatur angeführten, zählen unter anderem *Number of employees*, *Workforce diversity*, *Cost per hire*, *Span of control*, *Absenteeism* und *Number of qualified candidates per position* (International Organization for Standardization 2018).

V. AUSWAHL EINES GEEIGNETEN VERGLEICHSMODELLS

Das folgende Kapitel beschäftigt sich damit, einen geeigneten Kriterienkatalog für den Vergleich der genannten Business-Intelligence-Lösungen zu erarbeiten. Ähnliche, teilweise weitaus umfangreichere Vergleiche wurden in der Vergangenheit bereits durchgeführt und zeichnen seit Jahren marktführende Unternehmen und deren Software aus. Der *Magic Quadrant* der Gartner Gruppe zählt zu einem der bekanntesten Vergleiche der Business-Intelligence-Branche. Dessen Ergebnisse können für diesen Vergleich allerdings aus zwei Gründen nur eingeschränkt herangezogen werden:

Erstens gibt die Gartner Gruppe zwar der Bewertung zugrunde liegende Vergleichskriterien an, es ist jedoch

nicht nachvollziehbar, mit welcher Gewichtung sie in die Beurteilung einfließen oder welche Subkriterien diese enthalten. Alle Kriterien werden mit lediglich einem Satz beschrieben. Bei der Gegenüberstellung der bewerteten Funktionalitäten fällt es schwer, von einer gleichen Gewichtung auszugehen. Nicht alle Anforderungen an eine Business-Intelligence-Software sind von gleicher Bedeutung für den Endnutzer (Lindemann 2009). Zudem bewertet die Gartner Gruppe auch Aspekte wie die Sicherheit der Software, welche im folgenden rein funktionalen Vergleich nicht abgedeckt werden.

Der zweite deutlich ausschlaggebendere Punkt bezieht sich auf das Anwendungsumfeld der Software. Während der *Magic Quadrant* von einem allgemeinen Einsatz in Unternehmen ausgeht, erfolgt die Bewertung im Rahmen dieses Vergleichs speziell angepasst an das Aufgabenumfeld im Bereich People Analytics. Die Herausforderungen, wie sie hier auftreten, weichen mitunter stark von denen anderer Unternehmensbereiche ab.

Aus den genannten Gründen wird für diese Untersuchung ein anderes Bewertungsmodell benötigt, das sich am Vergleich der Gartner Gruppe orientiert und diesen ausdifferenziert. Um die Auswahlkriterien unterschiedlich gewichten zu können, müssen sie zuerst klassifiziert werden. Das spiegelt den Bedarf nach bestimmten Funktionalitäten besser wieder. Zusätzlich entfällt so die Notwendigkeit einer neuen vollumfänglichen Evaluierung, da sich zukünftige Veränderungen besser berücksichtigen lassen. Ziel ist es, eine höchstmögliche Zufriedenheit unter Betrachtung des funktionalen Softwareumfangs, der visuellen Endergebnisse und dem damit verbundenen zeitlichen Aufwand zu erreichen. Für das Erstellen des Auswahlkatalogs wird deswegen auf die theoretischen Grundlagen des Kano-Modells zurückgegriffen.

Das Kano-Modell kommt ursprünglich aus dem Bereich der Produktentwicklung und dient der Identifikation von verschiedenen Produkteigenschaften und deren Kombination, die in ihrer Gesamtheit die Kundenzufriedenheit maximieren. Dabei wird zwischen Basis-, Leistungs- und Begeisterungsanforderungen unterschieden (Dalton 2019). Das Modell vertritt den Ansatz, dass der Erfüllungsgrad und das Vorhandensein von Kundenanforderungen unterschiedliche Effekte auf die Zufriedenheit mit dem Produkt haben. Basisanforderungen bezeichnen dabei grundlegende Funktionen, die zu Unzufriedenheit führen, sofern sie fehlen. Leistungsanforderungen können sich je nach Ausprägung positiv oder negativ auf die Zufriedenheit auswirken, während Begeisterungsfaktoren einen rein positiven Effekt haben (Hölzing 2008).

VI. ANWENDUNG DES KANO-MODELLS

Durch die Verwendung des Kano-Modells wird Anforderungen eine Gewichtung verliehen (Lindemann 2009), „[...] [so] können jene Produktkriterien identifiziert werden, welche den größten Einfluss auf die Zufriedenheit des Kunden haben“ (Sauerwein 2000). Auch wenn das Kano-Modell üblicherweise vor der Produktfertigstellung zum Einsatz kommt, kann es auch abgewandelt für den

Vergleich von Produkten herangezogen werden. Business-Intelligence-Anwendungen sind keine Individualsoftware und in ihrem funktionalen Umfang nicht unternehmensspezifisch anpassbar. So müssen Unternehmen eine geeignete Anwendung aus einem Softwarepool wählen. Je mehr eine Software die individuellen Anforderungen der Kunden erfüllt, desto geeigneter ist sie laut dem Kano-Modell (Roth und Bernardy 2019). Basis-, Leistungs- und Begeisterungsanforderungen unterscheiden sich zusätzlich je nach Nutzerkreis, folglich ist es möglich, den Vergleich speziell an die Bedürfnisse von People Analytics anzupassen (Sauerwein 2000). Das Kano-Modell liefert so Hilfestellung beim Abwägen von Produkteigenschaften und -funktionen. Es unterstützt den Prozess, Kriterien mit dem größtmöglichen Einfluss auf die Kundenzufriedenheit zu identifizieren, auch wenn diese die festgelegten Anforderungen nicht vollkommen erfüllen. Diese Herangehensweise erleichtert die Entscheidung für eine geeignete Business-Intelligence-Software für den Einsatz im Bereich People Analytics bei Vitesco Technologies.

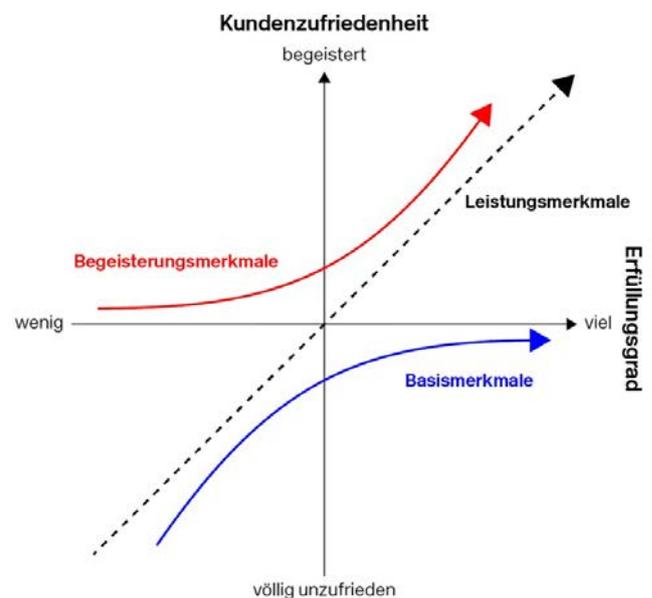


Abbildung 1. Kano-Modell (Bearbeitet nach (Jesse 2020))

A. Bestimmung und Klassifizierung der Anforderungen

Der erste Schritt im Vorgehen nach dem Kano-Modell beinhaltet das Erheben verschiedener Anforderungen. Diese werden normalerweise mithilfe eines Kano-Fragebogens klassifiziert (Sternad und Mödritscher 2018). Aufgrund der geringen Mitarbeiterzahl im People-Analytics-Team ergibt die Erstellung eines quantitativ ausgelegten Fragebogens keinen Sinn, da dieser kaum statistische Aussagekraft besitzt. Deshalb wurden Wünsche und Anforderungen an eine Business-Intelligence-Software im Team diskutiert und zusammen entsprechende Kano-Kriterien erarbeitet. Es folgen Definitionen aller als relevant erachteten Kriterien samt Umfang und Zuordnung zur entsprechenden Kano-Merkmalsskategorie:

1) **Basismerkmale:**

Datenzugriff und Anbindung von Quellsystemen

Die Vielfalt an verschiedenartigen Datenquellsystemen mit besonderem Fokus auf der Nutzung und Unterstützung von Clouddatenquellen.

Hier wird untersucht, wie gut die zugrunde liegenden Datentypen und Beziehungen im Einbindeprozess erkannt und relational sowie hierarchisch angereichert werden können. Ferner erfolgt eine Untersuchung auf Fehlertoleranz des Quellenbindungsprozesses und daraus resultierendem Nachbereitungsaufwand.

Graphical-User-Interface und SaaS-Weboberfläche

Der funktionale Umfang, Gestaltungsmöglichkeiten und intuitive Nutzung der Anwendung aus Entwickler- und Endanwendersicht.

Schwerpunkt für die Beurteilung aus Endnutzersicht ist der klare und verständliche Aufbau der Anwendung, also wie schulungsintensiv die Verwendung der Software für Gelegenheitsnutzer ist. Aus Entwicklerperspektive fließen Erstellungs- und Verwaltungsaufwand neuer Berichte, der funktionale Umfang sowie graphische Anpassungsmöglichkeiten in den Vergleich ein.

Integrierte ELT-Prozesse und KPI-Berechnungen

Der Funktionsumfang von ELT-Prozessen und individueller KPI-Berechnungen.

Dieses Bewertungskriterium deckt nachträgliche Anpassungsmöglichkeiten vorhandener Datenquellen samt deren Datentypen und -strukturen sowie die Erfassung der Bearbeitungsschritte ab. Korrektur- und Fehleridentifikationsprozesse sind ein großer Bestandteil, insbesondere der Korrekturaufwand und die Nachhaltigkeit vorgenommener Korrekturen. Der Vergleich berücksichtigt ebenfalls die Existenz von KPI-Berechnungsvorlagen und deren Fehleranfälligkeit bei Unregelmäßigkeiten in der Datenbasis. Ferner fließen Qualität und Quantität der ELT-Prozess- und KPI-Funktionsdokumentation hier mit ein.

2) **Leistungsmerkmale:**

Leistung der Analytics Engine

Die verfügbare Leistung der Analytics Engine, insbesondere in der Cloud.

Schwerpunkt sind analytische Modelle darunter integrierte Data-Mining-Prozesse oder prädiktive statistische Verfahren. Faktoren wie eine intelligente und korrekte Erkennung von Datentypen, Beziehungen und Mustern innerhalb und zwischen Quellen fließen in dieses Bewertungskriterium ein. Die Komplexität dieser Prozesse und Verfahren wird primär aus Entwicklersicht betrachtet, in erster Linie der benötigte Einarbeitungsaufwand und Programmierkenntnisse.

Integriertes Nutzer- und Rechtemanagement

Verwaltung von Nutzern, deren Nutzergruppen und Zugriffsrollen.

Bewertet wird der Arbeitsaufwand bei der Erstellung von Nutzerkonten und -rollen sowie der Berechtigungsvergabe. Der Fokus liegt auf dem langfristigen Änderungsaufwand und Verwaltungsaufwand, vor allem bei individuellen und umfangreichen Freigaben. Wichtig ist ebenfalls die Existenz von Single-Sign-On-Verfahren, also der Authentifizierung durch bereits vorhandene Konten.

3) **Beigeisterungsmerkmale:**

Self-Service-Analytics-Unterstützung

Die Möglichkeit, direkt und eigenständig auf verfügbare Datenquellen und analytische Funktionen zuzugreifen.

Hierzu zählen Anpassungen existierender Berichte, einfache, vom Nutzer anwendbare OLAP-Funktionen (Slicing, Dicing und Drill Downs), das Setzen von Filtern auf vorhandene Daten, die visuelle Veränderung von Diagrammen und die Erstellung eigener lokaler Dashboards. Diese Kriterien werden vorwiegend unter Berücksichtigung des funktionalen Umfangs, der intuitiven Nutzung und des Schulungsaufwands betrachtet.

Ad-Hoc-Reporting-Funktionalitäten

Das kurzfristige Erstellen von Berichten basierend auf einer weniger umfangreicheren, spezifischen Datenbasis.

Das Erstellen der benötigten Datenbasis sollte einfach mit geringem zeitlichem Aufwand und mit möglichst wenig personellen Ressourcen realisierbar sein. Folglich wird hier die Komplexität der Ad-Hoc-Berichterstellung und die Wiederverwendbarkeit vorhandener Datenmodelle – die Umsetzung des Model-View-Paradigmas – bewertet.

B. Umwandlung der Kano-Anforderungen in einen Kriterienkatalog

Aufgrund des intern festgelegten Kriterienkatalogs kann die Gewichtung der verschiedenen Kano-Kategorien nicht über die in der Literatur genannten statistischen Verfahren erfolgen (Vgl. hierzu Hölzing 2008; Sauerwein 2000). Für den vorliegenden Anwendungsfall der Evaluation zweier ähnlicher Produkte kann die Gewichtung der Überkategorien untergeordnet betrachtet werden, wenn diese bereits durch die Einzelbewertung der Kano-Merkmale garantiert wird. Um das sicherzustellen, erhält jedes Merkmal eine getrennte Bewertung zwischen -10 und +10 Bewertungseinheiten. Eine höhere Punktzahl kennzeichnet eine starke Ausprägung des Erfüllungsgrads und zeigt somit eine größere Zufriedenheit an.

Leistungsmerkmale tragen je nach Prägnanz zu einem linearen Punktezuwachs oder -verlust bei. Stattdessen nehmen die Punkte bei Begeisterungs- und Basismerkmalen in Richtung der Skalenenden stark zu oder ab. Geringfügige Unterschiede bestehen lediglich in den Grenzbereichen letzterer. Die genaue Verteilung kann Abbildung 2 entnommen werden. Abschließend werden Punktedurchschnitte der Kano-Kategorien und deren Mittelwert berechnet, dieser bildet das Endergebnis.

Anforderungstyp	Erwartungen übertroffen	Erwartungen erfüllt	Erwartungen teilw. erfüllt	Erwartungen kaum erfüllt	Erwartungen nicht erfüllt
Basisanforderung	0	0	-5	-7	-10
Leistungsanforderung	10	5	0	-5	-10
Begeisterungsanforderung	10	8	5	2	0

Abbildung 2. Kano-Kriterien Punkteverteilung (Eigene Darstellung)

C. Funktionale Analyse

1) Analyse Power BI:

Datenzugriff und Anbindung von Quellsystemen:

Power BI unterstützt ein breites Spektrum an Datenquellen, darunter *Webqueries*, lokale Ordnerstrukturen und unstrukturierte Daten wie PDFs. Die starke Integration in die Microsoft-Produktfamilie sorgt für einfache und schnelle Anbindungen zu SharePoints und der Power- sowie Azure-Plattform. Firmeninterne On-Premise-Systeme benötigen zur Kommunikation aus Sicherheitsgründen ein Power-BI-Gateway (Microsoft Corporation 2022).

Da die Software lediglich Daten abgefragter Quellen zwischenspeichert, muss sie nach einer Aktualisierung alle zugrunde liegenden Daten erneut abrufen. Dadurch entstehen bei umfangreichen Datensätzen, wie sie im Personalmanagement vorhanden sind, auch nach kleinen, lokal eingrenzenden Änderungen längere Wartezeiten im Vergleich zur SAC. Führt ein Quellabruf zu einem Fehler, bedingt dies einen Abbruch der Aktualisierung. Eine Wiederherstellung der vorherigen Daten findet statt. Das Abfragen von Datenbanken findet per vollständigem Vorabdatenimport oder *DirectQuery* statt. Letztere verwendet die Anmeldeinformationen des aktuellen Nutzers für die Datenabfrage.

Datentypen werden automatisch erkannt und strukturelle Anpassungen sind zur Laufzeit möglich. Dies erhöht die Flexibilität nachträglicher Anpassungen, allerdings birgt die Kombination aus manuellen und automatischen Arbeitsschritten viele potenzielle Fehlerquellen. Inhaltliche Überprüfungen erfolgen nur bedingt. Resultierend entsteht ein erhöhter nachträglicher Korrekturaufwand.

Graphical-User-Interface und SaaS-Weboberfläche:

Power BI unterteilt die GUI in eine Entwickler- und eine Endanwenderapplikation. Lediglich Letztere ist als SaaS-Webanwendung umgesetzt und dient primär dem Teilen von Dashboards. Nur die Desktopanwendung ermöglicht das Bearbeiten von Dashboards und Berichten. In dieser Anwendung werden Datenverbindungen festgelegt, ETL-Prozesse vorgenommen und Dashboards erstellt. Dieser Softwareaufbau resultiert in einer Betriebssystem- und Ressourcenabhängigkeit.

Power BI überzeugt hier vor allem mit einer großen Auswahl an Diagrammen und graphischen Elementen. Diese können mit anderen Darstellungen und Filtern interagieren oder individuell von Interaktionen ausgeschlossen werden. Die Darstellung eines zeitlichen Verlaufs bleibt so trotz eines monatlichen Filters weiterhin unbeeinflusst. Fast jedes Diagramm unterstützt mehrdimensionale Darstellungsebenen und OLAP-Funktionen benötigen keine strikt hinterlegten Relationen in der Datenbasis.

Funktional programmierbare Schaltflächen unterstützen den Wechsel zwischen Berichtsseiten und Lesezeichen, können Diagramme jedoch nicht inhaltlich verändern. Viele dieser Elemente führen zu einer hohen Anzahl an zu verwaltenden Lesezeichen und somit zusätzlichem Verwaltungsaufwand.

Integrierte ELT-Prozesse und KPI-Berechnungen: Da die Anwendung Daten lediglich aus Quellen abrufen und nicht erneut abspeichert, durchlaufen diese kein festgelegtes Datenmodell und folglich weniger Korrektheitsprüfungen. Daraus resultierend gelangen viele nicht plausible Werte in die Rohdatenmenge. Korrekturen und Überprüfungen obliegen dem Ersteller. Die Datenbasis kann zwar mehr Fehler enthalten, ist aber gleichzeitig robuster und toleranter gegenüber Unregelmäßigkeiten.

Für Datenänderungen und -erweiterungen stehen unterschiedliche Funktionen zur Verfügung. Durch eine Ähnlichkeit zu Excel wirkt das Anwendungslayout an vielen Stellen bekannt. Benutzerdefinierte Spalten können mithilfe der Programmiersprache F# erstellt werden. Diese unterstützt auch komplexe Berechnungen auf Indexbasis, ist aber stark durch die lokale Rechenleistung eingeschränkt. Getätigte Verarbeitungsschritte werden spaltenübergreifend dokumentiert und erscheinen deswegen oft unübersichtlich. Identische Spaltenanpassungen werden nicht zusammengeführt, stattdessen erfolgt eine erneute Transformation der bereits modifizierten Spalte. Viele Bearbeitungsschritte führen zu unübersichtlichen Extract-Load-Transform&Transform-again-Prozessen. Das Entfernen vorangegangener Arbeitsschritte führt besonders bei Änderung der Spaltenbezeichnungen zu aufwändigen Fehlerkorrekturen.

Während die Anpassung der Datenquellen unter Verwendung einer Programmiersprache erfolgt, werden KPIs mithilfe von syntaktisch und semantisch anders aufgebauten DAX-Formeln erstellt. Deren Aufbau ähnelt ebenfalls Excel-Funktionen und setzt so weniger Kenntnisse in objektorientierten und imperativen Programmiersprachen voraus. Mitarbeiter benötigen folglich weniger Einarbeitungsaufwand, um mit diesen arbeiten zu können. Die verständlicheren Formeln verleiten allerdings zu weiteren Anpassungen abseits der Datenquellen. Sie sollten beachtet werden, da die Fehlerbehebung nur in den betroffenen Berichten, aber nicht nachhaltig in den Quellsystemen erfolgt.

Sowohl für F# als auch DAX-Formeln existieren ausreichend Schulungsmaterialien im Internet. Ein Großteil der auftretenden Fehler ist ohne Experten mithilfe von Lernvideos und Blogbeiträgen schnell und selbstständig lösbar.

Leistung der Analytics Engine: Die Analytics Engine ist in vielerlei Hinsicht durch die lokale Anwendung und damit einhergehenden Ressourcenlimitierungen eingeschränkt. Dennoch bietet Power BI einige Data-Mining-Werkzeuge. Mit dem *Q&A-Feature* können unter Verwendung natürlicher Sprache Fragen gestellt werden. Die Analytics Engine wertet diese aus und liefert das entsprechende Ergebnis in Form eines passenden Diagramms. Die Kennzahlen und Filterattribute müssen dabei mit den Spaltenbezeichnungen des Datenmodells übereinstimmen. Eine manuelle Erweiterung der Sprach- und Wortvielfalt ist möglich, erscheint aber sehr arbeits- und zeitintensiv.

Beziehungen zwischen Quellen und vorhandene Datentypen werden automatisch erkannt und sind nachträglich in ihrer Relationsart und Kommunikationsrichtung anpassbar. Die selbstständige Erkennung identifiziert bei ähnlichen Spaltenwerten oft falsche Beziehungen. Zwischen zwei Datenquellen kann stets nur eine Verbindung existieren, weitere werden von der Software erkannt, aber als inaktiv hinterlegt. Spalten innerhalb einer Quelle relational zu verknüpfen, ist nicht möglich. In einer speziellen Modellansicht lassen sich alle Beziehungen graphisch aufbereitet einsehen und bearbeiten.

Auf Zeitreihen basierende Diagramme können um Prognosen erweitert werden. Diese sind durch Variation der Länge, Festlegung des Konfidenzintervalls und der Hinterlegung von Saisonalität anpassbar. Zudem können Perioden gezielt von der Berechnung ausgeschlossen werden. Auch die Einarbeitung in hier genannte Funktionen fällt dank der umfangreichen Online-Lernmaterialien gering aus.

Integriertes Nutzer- und Rechtemanagement: Die Erstellung von Nutzerrollen und deren Berechtigungen geschieht über die Desktopsoftware, die Nutzerzuordnung allerdings über die SaaS-Anwendung. Das Sicherheitskonzept basiert auf einer Row-Level-Security und der Kopplung zwischen der Rolle und dem Microsoft-Benutzerprinzipalnamen, einem Authentifizierungsattribut. Einschränkungen sind quellenunabhängig auf Basis zugrunde liegender Zeilen hinterlegt.

Spezielle Freigaben können basierend auf dem Benutzerprinzipalnamen in den Rollen festgehalten werden. Nutzer müssen letzteren zusätzlich noch manuell in der Webanwendung zugeordnet werden. Dafür laufen Authentifizierung und der Abruf zugewiesener Rollen über das Microsoftkonto. Ein zusätzliches Anlegen von einzelnen Nutzerkonten entfällt so.

Einschränkungen erfolgen separat in jeder Tabelle. Bestehen viele verschiedene Rollen oder Nutzer mit abweichenden individuellen Zugriffsberechtigungen, entsteht ein großer Verwaltungsaufwand. Im Bereich People Analytics existieren nur wenige Personen, die über einen vollumfänglichen Lesezugriff verfügen. Normale Nutzerrollen besitzen meist mindestens eindimensionale Zugriffsbeschränkungen, was in Zusammenhang mit der aufwendigen Nutzerpflege zu einem hohen Arbeitsaufwand führt.

Self-Service-Analytics-Unterstützung: Power BI verfügt über kaum spezielle Self-Service-Analytics-Funktionen. Nutzer können mit Diagrammen interagieren, abgebildete Daten ihren Rollen entsprechend einsehen und extrahieren sowie Filter setzen. Das Erstellen eigener Dashboards auf Basis eines bereits existierenden Datenmodells ist allerdings nicht möglich, da dieses fest mit dem Dashboard verknüpft ist. Verfügen Data Warehouses oder ähnliche Live-Datenverbindungen über alle Datensätze des Datenmodells, sollte aufgrund des indirekten Quellabrufs trotzdem eine Überprüfung der Metadaten erfolgen. Die Power-BI-Desktopanwendung ist grundsätzlich kostenlos verfügbar

und ermöglicht jedem das Erstellen von Dashboards, die Modellierung der Datenbasis ist für Gelegenheitsnutzer jedoch zu zeitaufwendig. Das benötigte Gesamtverständnis führt dazu, dass individuelle Wünsche über Filter und zusätzliche Diagramme in einem zentralen oder mehreren spezifischen Dashboards umgesetzt werden müssen. Das erhöht den Verwaltungsaufwand, garantiert aber einheitliche Datenqualität und unterstützt die Single-Source-of-Truth-Berichterstattung von People Analytics.

Die Q&A-Funktionalitäten ermöglichen dennoch eingeschränkte Selbstevaluierungen mittels eigener Darstellungen. Diese sind allerdings nur temporär verfügbar und deswegen lediglich für einmalige Abfragen geeignet. Schließt man die Anwendung, werden sie zurückgesetzt.

Ad-Hoc-Reporting-Funktionalitäten: Aufgrund des indirekten Datenzugriffs und der berichtsgebundenen Abhängigkeit der ELT-Prozesse eignet sich Power BI nur bedingt zum Erstellen von Ad-Hoc-Berichten. Power BI übernimmt beim Kopieren vorhandener Berichte einerseits getätigte Veränderungen im Datenset und berechnete KPIs. Das führt andererseits zu einem hohen Aufwand bei nötigen Anpassungen der Datenquellen, primär wegen der hohen Fehleranfälligkeit bei Modifikationen der ELT-Bearbeitungsschritte.

Liegen keine entsprechende Datenbasis oder ähnliche Berichte vor, können Daten mithilfe der Bearbeitungsfunktionen jedoch einfach und schnell modifiziert werden. Bei kleineren Datensätzen besteht so keine Notwendigkeit einer Vorabanpassung durch zusätzliche vorhergehende Datenbearbeitungsprozesse oder weitere Softwareanwendungen.

Dazugehörige Datensätze der Dashboards sind in der Webanwendung gespeichert. Viele Berichte führen hier zu einer unübersichtlichen Struktur, deren Verwaltung zusätzliche personelle Kapazitäten erfordert, insbesondere wenn regelmäßige Aktualisierungen und Überprüfungen notwendig sind.

2) Analyse SAP Analytics Cloud:

Datenzugriff und Anbindung von Quellsystemen: Die SAC unterstützt drei verschiedene Datenquellen: einzelne Dateien, vorab erstellte Modelle und weitere, überwiegend auf SAP basierende Clouddatenquellen. Um auch bei der Verwendung von lokalen Dateien die Vorteile der Cloud zu nutzen, speichert die SAC diese vor der Einbindung im mitgelieferten Clouddatenverzeichnis ab. Ein direkter Abruf von REST-APIs und Webqueries ist nicht möglich.

Bei der Erstellung von Datenmodellen erfolgt eine Unterteilung der Spalten in Dimensionen und Kennzahlen. Dimensionen ermöglichen das Aufteilen und Spezifizieren von Kennzahlen basierend auf ausgewählten Eigenschaften und Attributen. Dabei wird grundsätzlich davon ausgegangen, dass Daten entweder eine Kennzahl oder eine Dimension darstellen, was speziell im Bereich People Analytics nicht immer der Fall ist. Die für Berechnungen relevanten Dimensionen erfordern eine nachträgliche Konvertierung in eine Kennzahl.

Nützlich hingegen ist das Verknüpfen von Dimensionen innerhalb einer Quelle. Spalten können hierarchisch, relational oder erweiternd als Beschreibung miteinander verbunden werden. Im Dashboard sind diese dann zu einer Dimension zusammengefasst. Bei komplexen personenspezifischen Hierarchien zeigt die SAC allerdings Schwächen in der technischen Umsetzung. Spezialfälle erschweren die Abbildung von Hierarchien. Ein Beispiel im Bereich People Analytics sind temporär im Ausland tätige Mitarbeiter, welche geografisch ihrem Zielland, aber gleichzeitig weiterhin ihrer heimischen Unternehmensgesellschaft zugeordnet sind. Dieses Problem wirkt sich auch auf die Übersichtlichkeit der graphischen Darstellungen aus.

Derart strikte technische Einschränkungen existieren an mehreren Stellen. Sicherheitsvorkehrungen, deren Existenz generell berechtigt ist, führen zu einer arbeitsintensiven Nachbereitung, vor allem bei regelmäßigen Anpassungen. Dafür dienen sie als weitere Sicherheitsvorkehrung in Bezug auf Datenqualität und verhindern das Hinterlegen falscher Relationen. Bei einer entsprechend qualitativen Datenbasis und gefestigten Modellen sind die Vorteile einer cloudbasierten Datenhaltung deutlich spürbar.

Graphical-User-Interface und SaaS-Weboberfläche:

Die Webanwendung dient als Entwicklungsumgebung und Plattform für Endanwender. Die Erstellung und Bearbeitung von Berichten und Dashboards erfolgt wie üblich bei SaaS-Anwendungen im Webbrowser. Die SAC ist somit komplett betriebssystemunabhängig nutzbar. Entwickler können neben *Stories*, hierunter fallen einzelne Berichte oder Dashboards, auch *Analytic Applications* erstellen. Letztere unterstützen die Gestaltung interaktiver Berichte durch die Verwendung von Skripten.

Die Auswahl an Diagrammen, Filtern und Funktionen fällt hier, wie auch bei den ELT-Prozessen, vergleichbar gering, dennoch ausreichend aus. Komplexe Designwünsche können einerseits mit der SAC nur schwer umgesetzt werden, denn Formen zur graphischen Aufbereitung sind kaum vorhanden und Anpassungsmöglichkeiten, die über eine Änderung der Füllfarbe hinausgehen, fehlen. Eine ähnliche Problematik tritt bei komplexeren, mehrdimensionalen Diagrammen auf. Die wenigen Anpassungsmöglichkeiten führen andererseits zu übersichtlich gestalteten und leicht verständlichen Berichten.

Die SAC unterscheidet zwischen Seiten- oder Storyfiltern. Letztere besitzen einen eigenen Abschnitt in der Anwendung und beziehen sich auf alle Berichtsseiten. Während die Interaktionen zwischen Diagrammen anpassbar sind, können einzelne Darstellungen nicht von der Interaktion der Storyfilter ausgeschlossen werden. Eine Kombination aus Zeitreihen und monatlichen Kennzahlen benötigt so umfangreiche Anpassungen und Programmierkenntnisse.

Spezifische Filter und Darstellungsmethoden können unter der Verwendung von *Analytic Applications* und freiprogrammierbarer Skripte realisiert werden. Die SAC ermöglicht hier das Erstellen einseitiger Webapplikatio-

nen, die eine funktionale Erweiterung der Nutzereingaben und Filter unterstützen. Hierfür kommt die Programmiersprache JavaScript zum Einsatz (SAP SE 2020b). Diese Skripte greifen direkt auf einzelne Diagramme zu und erlauben dynamische Anpassung zur Laufzeit sowie Abfragen der Diagramminhalte. Mittels einer integrierten API unterstützt die SAC auch die Vernetzung mehrerer Anwendungen und Berichte.

Integrierte ELT-Prozesse und KPI-Berechnungen: Vor dem endgültigen Erstellen eines Datenmodells können Spalten in einer Vorschau verwaltet und hinzugefügt sowie Daten manipuliert und Fehler behoben werden. Im Vergleich zu Power BI gibt es hier deutlich weniger Funktionen, trotzdem werden (fast) alle benötigten Anpassungsarten unterstützt.

Programmierähnliche Funktionen ermöglichen die Ergänzung von benutzerdefinierten Spalten und eine repräsentativ gestaltete Stichprobe begünstigt Echtzeitüberprüfungen. Diese erleichtert die Validierung berechneter Werte. Berechnete Spalten können beim Hinzufügen weiterer Dateien allerdings zu Problemen führen, denn nach der Modellerstellung erfolgt bei weiteren Importen keine Unterscheidung zwischen im Rohdatensatz existierenden und manuell berechneten Spalten.

Ein spaltenspezifisches Transformationsprotokoll hält die Bearbeitungsschritte fest. Änderungen einzelner Spalten führen so nicht zur Notwendigkeit einer erneuten Gesamtüberprüfung. Die Ressourcenskalierung der Cloud garantiert zudem schnelle und konstante Berechnungen.

Die Fehlerkorrektur unterstützt nur das Ersetzen oder Löschen der fehlerhaften Werte, sie können nicht ignoriert werden. Das stellt People Analytics vor eine große Herausforderung, denn inkorrekte Daten erfordern vollumfängliche Vorabkorrekturen oder es muss stattdessen auf Funktionalität verzichtet werden. Alle Spalten besitzen bereits während der Bearbeitung eine graphische Übersicht der vorhandenen Werte. Fehlerhafte Werte werden zusammen mit einer kurzen Fehlermeldung angezeigt. Die SAC führt zusätzlich Logikprüfungen durch, etwa ob ein Datum plausible Kalenderjahre über- oder unterschreitet.

Individuelle KPIs können später im Bericht erstellt werden. Dazu steht dieselbe, funktional weniger umfangreiche Eingabeaufforderung wie im Datenmodell zur Verfügung. Vorlagen für die Berechnung von Abweichungen, Aggregationen oder Einschränkungen vorhandener Kennzahlen existieren bereits.

Die Dokumentation der Funktionen ist sehr kurz gehalten und deswegen nur bedingt hilfreich. Die integrierte Eingabehilfe fordert bei Fehlern nur dazu auf Funktionen umzuschreiben. Genaue Fehlermeldungen oder ein Verweis auf das *SAP Analytics Cloud Help Portal* fehlen.

Leistung der Analytics Engine: Die Leistung der SAC Analytics Engine ist deutlich spürbar, so sind Berichte und Daten schnell geladen und längere Wartezeiten treten dank der dynamischen Ressourcenverteilung nicht

auf. Beziehungen zwischen Quellen werden nicht automatisch erkannt, eine Verlinkung dieser ist aber nachträglich möglich. Die richtige Erkennung der Datentypen ist durch die festgelegten Kennzahlen sichergestellt.

Fast jede Darstellungsform verfügt über Data-Mining-Funktionalitäten, sogenannte *Smart Insights*. Die Freischaltung dieser erfolgt individuell für jedes Diagramm. *Smart Insights* liefern Zusammenhänge zwischen den abgebildeten Kennzahlen und den verfügbaren Dimensionen.

Linien-, Flächendiagramme und Zeitreihen können durch Prognosen erweitert werden. Die SAC unterscheidet hier zwischen einer automatischen und einer erweiterten Prognose. Letztere ist durch lineare Regression, dreifache exponentielle Glättung oder eigens festgelegte Werte spezifizierbar.

Der Einarbeitungsaufwand steht hier in starkem Zusammenhang mit den Datenmodellen, ist aber aufgrund zahlreicher Skript-Funktionalitäten und der damit verbundenen Programmiersprache als hoch einzustufen.

Integriertes Nutzer- und Rechtemanagement: Gruppenspezifische oder individuelle Berechtigungen werden durch Nutzer- und Gruppenrollen vergeben. Die Verwaltung und Zugrifferteilung obliegt den Systemadministratoren und erfolgt rein über die SaaS-Applikation. Freigaben umfassen *Stories*, *Analytic Applications*, Datenmodelle und SAC-Orderstrukturen.

Bei der Verwendung eines SAP-Data-Warehouses ist eine Übernahme der allgemeinen SAP-Benutzerfreigaben möglich. Nutzer besitzen so in der SAC Zugriff auf alle Daten, die ihnen auch sonst in den SAP-Systemen zur Verfügung stehen. Durch die Wiederverwendung von Freigaben entfallen redundante Arbeiten wie das Erstellen und Verwalten von Nutzerrollen, was besonders bei internen Auditierungen Zeit spart. Existiert eine derartige Business-Intelligence-Struktur, werden kaum Kapazitäten für die Nutzerverwaltung benötigt.

Alle Nutzer einer Rolle oder eines Teams können zusätzlich als Datei exportiert oder importiert werden. Das Verwalten und Prüfen von Rollen ist besonders unter dem Aspekt der Wiederverwendbarkeit positiv zu bewerten.

Single-Sign-On-Verfahren existieren und können durch jeden Identity Provider, der das Authentifizierungsframework SAML 2.0 unterstützt, erfolgen (SAP SE 2020a). Eine Authentifizierung über das Microsoft-Firmenkonto ist möglich.

Self-Service-Analytics-Unterstützung: Die SAC überzeugt hier besonders mit der Möglichkeit, eigene Dashboards anzulegen. Das Erstellen einer korrekten und geeigneten Datenbasis entfällt bei der Verwendung von SAP-Data-Warehouses. Durch die Übernahme der Nutzerfreigaben verfügt dort jeder Endnutzer über alle ihm freigegebenen Daten. Das fördert resultierend Self-Driven-Analytics-Ansätze.

Die OLAP-Funktionen sind umfangreich und übersichtlich gestaltet. Durch die Auswahl, das Hervorheben oder Exkludieren einzelner Datenpunkte

besteht die Möglichkeit spezifische Teildatensätze zu bilden. Die entstehenden Subdatenmengen können unter Verwendung der OLAP-Funktionen weiterbearbeitet werden. Drilldowns sind jedoch nur bei zuvor als hierarchisch hinterlegten Datenkorrelationen möglich. Das zugrunde liegende Datenmodell kann demzufolge OLAP-Funktionen stark einschränken.

Diagramme unterstützen zudem ein interaktives Setzen von Filtern und zeigen diese auch an. Sie können direkt über die betroffenen Diagramme zurückgesetzt werden. Der Nutzer kann selbst zusätzliche zeitspezifische Filter wie einen direkten Vergleich zu Vorperioden setzen. Des Weiteren existieren Verweise auf die verwendeten Datenquellen und auch ein numerischer Export der visualisierten Daten ist möglich.

Durch mangelnde, teils unverständliche Erklärungen und Dokumentationen benötigen Endanwender mit großer Wahrscheinlichkeit zeitintensive Schulungen. Diese erwartet somit eine Self-Service-Analytics-Plattform mit ausbleibenden Self-Learning-Materialien.

Ad-Hoc-Reporting-Funktionalitäten: Existieren Datenbasen oder Live-Verbindungen kann auf diese entsprechend der eigenen Freigaben zurückgegriffen werden. Nur KPIs und Berechnungen müssen neu erstellt werden, da diese direkt in den Berichten hinterlegt sind. Die strikten Vorgaben der Datenmodelle führen zu einer hohen Wiederverwendbarkeit – wie sie das Model-View-Paradigma vorsieht – und unterstützen damit maßgeblich beim Erstellen von Ad-Hoc-Berichten. Benötigen Berichte allerdings neue Daten oder reichen die existierenden nicht aus, entsteht ein hoher Erstellung- und Änderungsaufwand. Gleiches gilt für die Skripte der *Analytic Applications*, hier werden für einfache Anpassungen bereits spezielle Fachkenntnisse benötigt. Änderungen sollten ausführlich getestet werden.

Liegt eine passende Abfrage vor, bietet die SAC zusätzlich einen *Data Analyzer* zur Überprüfung von Ad-Hoc-Abfragen. Dieser kann Kennzahlen und Dimensionen in einer Tabelle darstellen und bietet OLAP-ähnliche Funktionen.

Resultierend aus den umfangreichen Self-Service-Analytics-Funktionen ergibt sich die Notwendigkeit weiterer Überprüfungen, etwa ob Nutzer von Ad-Hoc-Berichten profitieren oder lieber mit eigenen Berichten arbeiten. Um zusätzliche Arbeit zu vermeiden, sollte die Existenz dieser klar kommuniziert, regelmäßig auf Redundanzen geprüft und Zugriffsstatistiken erhoben werden. Zu viele individuelle Anpassungen und Darstellungen führen zu einem Verlust an Datenqualität und dem Vertrauen in die Datenbasis (Gole und Shiralkar 2020; Hilgefort 2019).

VII. ERGEBNISSE DES VERGLEICHS

Der durchgeführte Vergleich favorisiert aus funktionaler Sicht die Verwendung der SAC für den Unternehmensbereich People Analytics bei Vitesco Technologies. Während die SAC Power BI in den Basismerkmalen leicht unterliegt, überzeugt sie vor allem bei den Leistungs- und

Begeisterungsmerkmalen. Eine Visualisierung der Endergebnisse ist Abbildung drei zu entnehmen.¹

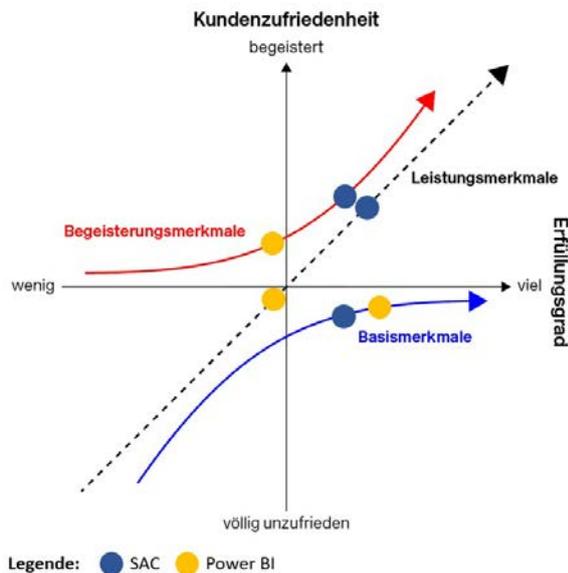


Abbildung 3. Visualisierung der Endergebnisse (Bearbeitet nach (Jesse 2020))

Power BI schafft es durch eine umfangreiche Auswahl an Datenquellen, vielfältige Gestaltungsmöglichkeiten und klar strukturierte umfangreiche ELT-Prozesse zu überzeugen. Das vorhandene Schulungsmaterial und eine große Power-BI-Community erleichtern das Arbeiten mit der Software, besonders beim Identifizieren und Beheben von Verarbeitungsfehlern. Durch die lokale Anwendung ist die verfügbare Leistung jedoch stark eingeschränkt und aufgrund mangelnder Ressourcen den hohen Anforderungen von People Analytics bei Vitesco Technologies nicht immer gewachsen. Erstellte Datenmodelle können nur schwer wiederverwendet werden, was das Erstellen von eigenen Dashboards und Ad-Hoc-Berichten stark einschränkt. Die vielen Nutzerrollen und deren Restriktionen führen zusätzlich zu einer arbeitsintensiven und fehleranfälligen Zugangsverwaltung.

Aufgrund der mangelnden Self-Service-Analytics-Funktionen und der eingeschränkten Leistung eignet sich Power BI im People-Analytics-Umfeld eher als zentrales Reporting-Instrument. Eine Erweiterung der Business-Intelligence-Anwendung durch zusätzliche Softwarekomponenten, die die anspruchsvollen Berechnungen übernehmen und so den Mangel an technischen Ressourcen ausgleichen, könnte der fehlenden Leistung teilweise entgegenwirken. Hier eignen sich etwa die Programmiersprache R oder die Datenanalysesoftware KNIME. Ob diese zusätzlichen Prozessschritte langfristig Arbeitsaufwand reduzieren und wirtschaftliche Profitabilität sichern, überschreitet den

¹Für detailliertere tabellarische Ergebnisse wenden Sie sich bitte an Fabian Engl bzw. Philipp Trubjansky unter den angegebenen Kontaktdaten

Umfang des Vergleichs und sollte als Bestandteil weiterer Untersuchungen genauer evaluiert werden.

Die SAC überzeugt vor allem durch die immer konstant verfügbare Leistung der Cloud. Durch die Erstellung von cloudbasierten Datenmodellen profitieren selbst lokale Dateien von den dynamischen Rechenressourcen. Mit der SaaS-Anwendung existiert eine zentrale Plattform für Entwickler und Endanwender, diese ist funktional allerdings weniger umfangreich. Konstante und feste Datenmodelle garantieren fehlerresistente Berechnungen, schränken aber Nutzer in der Datenmodellierung und -anpassung ein. Die strengen Restriktionen und Vorgaben treten besonders bei dimensionaler und relationaler Datenanreicherung auf. Dafür ermöglicht die Wiederverwendung von Datenquellen und Benutzerfreigaben ein einfaches Erstellen neuer Dashboards und Ad-Hoc-Berichte. Während die SAC Power BI auch in der graphischen Gestaltung deutlich unterliegt, ermöglichen eingebettete Funktionen objektorientierte sowie dynamische Zugriffe und Anpassungen aller Objekte. Die technisch umfangreichen und innovativen Funktionen der SAC leiden allerdings unter mangelhafter und teilweise fehlender Dokumentation sowie wenigen sonstigen Hilfestellungen. Resultierend daraus müssen Fehler und Funktionen eigenständig entdeckt und zeitintensiv korrigiert werden.

Die SAC ermöglicht Nutzern das eigenständige Auswerten und Visualisieren von Daten, folglich entfallen kundenspezifische Anpassungen zentraler Berichte und deren Datenquellen. Besteht allerdings keine Live-Datenverbindungen zu Data Warehouses oder SAP-Systemen, müssen fehlerfreie, logisch strukturierte und klar festgelegte Daten zugrunde liegen. Das stellt People Analytics bei Vitesco Technologies vor eine große Herausforderung. Wie sich die strikten technischen Vorgaben, der Schulungsaufwand für Endanwender und Entwickler sowie ein wahrscheinlicher Verlust der *Single Source of Truth* auf People Analytics bei Vitesco Technologies und deren innerbetriebliche Akzeptanz auswirkt, sollte vor einer endgültigen Entscheidung weiter untersucht werden.

VIII. FAZIT UND HANDLUNGSEMPFEHLUNG

Wie aus dem Vergleich erkenntlich wird, bevorzugt dieser aus funktionaler Sicht SAPs Analytics Cloud als zukünftige Business-Intelligence-Lösung für den Bereich People Analytics bei Vitesco Technologies. Das bedeutet allerdings nicht, dass die SAC die aktuelle Plattform Power BI vollumfänglich ersetzen kann und diese gänzlich ausgetauscht werden sollte. Zum Zeitpunkt des Vergleichs stellen vor allem fehlende Webqueries und überdurchschnittlich hohe Datenqualitätsanforderungen People Analytics bei Vitesco Technologies vor große Herausforderungen.

Solange die interne Berichterstattung weiterhin, wenn auch nur teilweise von Datenausgügen, CSV- und Excel-Dateien abhängig ist, benötigen die strikten Datenmodelle der SAC mehr Arbeitsaufwand als vergleichbare Modelle in Power BI. Besonders das Festlegen eines endgültigen People-Analytics-Datenmodells stellt unter Anbetracht des

kürzlichen Spin-offs von Vitesco Technologies ein großes Problem dar. Die sich immer noch im Wandel befindende HR-IT-Infrastruktur bringt sicherlich auch in Zukunft Änderungen der Reportingprozesse mit sich und kann deswegen noch nicht final betrachtet werden. Vor einem Wechsel zu einer neuen Business-Intelligence-Plattform müssen Datenqualität und -verfügbarkeit sichergestellt sein. Das bedeutet, alle für die Berichterstellung relevanten Daten sind über ein Data Warehouse verfügbar und grundlegende hierarchische und relationale Fehler wurden vorab in der Datenbasis behoben.

Um schon jetzt die innovativen Funktionen, insbesondere die Self-Service-Analytics-Eigenschaften der SAC, nutzen zu können, empfiehlt sich eine Übergangsphase, bis sowohl die HR- als auch die People-Analytics-Prozesslandschaft endgültig feststehen. Ein mögliches Szenario wäre die Migration von kleineren Berichten, welche bereits über eine endgültige und korrekte Datenbasis verfügen. Mit Pilotprojekten können Endnutzer an die Funktionalität und den neuen nutzerzentrierten Ansatz der SAC herangeführt werden. Hierfür eignen sich kleinere Berichte über Arbeitsunfälle oder vorhandene nationale Diversität. Mitarbeiter von People Analytics ermöglicht dieses Vorgehen zudem, erste Erfahrungen mit der SAC und deren *Analytic Applications* zu sammeln. Auf diese Weise machen sich sowohl Entwickler als auch Endnutzer langsam mit der Plattform vertraut. Weitere Dashboards sollten stufenweise folgen. Berichte, die auf Webinhalten und -abfragen basieren, müssen nach aktuellem Stand weiterhin mit Power BI umgesetzt werden.

Abschließend lässt sich sagen, dass die SAC den Bereich People Analytics um wertvolle Funktionalitäten und neue innovative Reportingansätze erweitern kann. Dafür müssen allerdings eine Reihe an Vorbedingungen erfüllt sein, diese erschweren einen sofortigen Umstieg und erfordern weitere Vorarbeit. Sobald alle benötigten Daten fehlerfrei über Data Warehouses oder direkte SAP-Systemanbindungen abrufbar sind, erfüllt sie alle Anforderungen für die interne Verwendung im People-Analytics-Umfeld. Durch den Fokus auf Live-Datenverbindungen und einem klar strukturierten Cloud-Business-Intelligence-Konzept erscheint ein Wechsel auf langfristige Sicht vielversprechend.

IX. ZUSAMMENFASSUNG

Aufgrund leistungsbedingter Einschränkungen durch die aktuelle Business-Intelligence-Software Power BI vergleicht die People-Analytics-Abteilung von Vitesco Technologies diese mit der Alternativsoftware SAP Analytics Cloud. Dafür wurden zunächst aktuelle Herausforderungen im People-Analytics-Umfeld identifiziert und basierend darauf Vergleichskriterien erarbeitet. Als Vergleichsmodell kommt das Kano-Modell zum Einsatz. Die durchgeführte Evaluation favorisiert aus funktionaler Sicht einen Umstieg auf die SAP Analytics Cloud, identifiziert allerdings eine Reihe an Herausforderungen, die einen sofortigen Wechsel einschränken. Zu diesen gehören sowohl die Verfügbarkeit als auch die Qualität der HR-Daten.

LITERATUR

- Angrave, D. u. a. (2016). "HR and analytics. Why HR is set to fail the big data challenge". In: *Human Resource Management Journal* 26.1, S. 1–11.
- Capgemini (Feb. 2021). *Studie IT-Trends 2021. IT ermöglicht Business trotz Kontaktbeschränkungen*. <https://de.statista.com/statistik/studie/id/87286/dokument/umfrage-zum-stand-der-it-in-unternehmen-2021/>. (besucht am 30.01.2022).
- Contact-Center-Network (2020). *Haben Sie für das Jahr 2020 Investitionen in Business Intelligence in Ihrem Contact Center geplant?* <https://de.statista.com/statistik/daten/studie/804998/umfrage/umfrage-zum-geplanteneinsatz-von-cloud-loesungen-im-contact-center/>. (besucht am 30.01.2022).
- Dalton, J. (2019). *Great Big Agile: An OS for Agile Leaders*. 1. Aufl. Springer Science + Business Media Finance Inc, S. 189.
- Devens, R. M. (1865). *Cyclopaedia of Commercial and Business Anecdotes*. 1. Aufl. D. Appleton Company, S. 210.
- Eckerson, W. W. (2009). *Performance management strategies. How to Create and Deploy Effective Metrics*. <https://mindsight.com.br/wp-content/uploads/2020/08/How-to-Create-and-Deploy-Effective-Metrics-by-Weyne-Eckerson.pdf>. (besucht am 30.01.2022).
- Filipowicz-Florczyk, A. (2014). "Key Performance Indicators of the HR Function". In: *Organizacja i Zarzadzanie* 28.4, S. 75–88.
- Gani, A. u. a. (Juni 2016). "A survey on indexing techniques for big data. Taxonomy and performance evaluation". In: *Knowledge and Information Systems* 30.3, S. 18–23.
- Gartner Inc. (Dezember 2021). *Gartner auf einen Blick*. <https://emtemp.gcom.cloud/ngw/globalassets/intl-de/ueber/dokumente/gartner-auf-einen-blick.pdf>. (besucht am 30.01.2022).
- Gluchowski, P. (Juni 2016). "Business Analytics – Grundlagen, Methoden und Einsatzpotenziale". In: *HMD Praxis der Wirtschaftsinformatik* 53.3, S. 273–286.
- Gluchowski, P., R. Gebriel und C. Dittmar (2008). *Management Support Systeme und Business Intelligence. Computergestützte Informationssysteme für Fach- und Führungskräfte*. 2. Aufl. Springer Verlag, S. 273.
- Gole, V. und S. Shiralkar (2020). *Empower Decision Makers with SAP Analytics Cloud. Modernize BI with SAP's Single Platform for Analytics*. 1. Aufl. Springer Science + Business Media Finance Inc, S. 24.
- Gurjar, Y. S. und V. S. Rathore (Jan. 2013). "Cloud Business Intelligence – Is What Business Need Today". In: *International Journal of Recent Technology and Engineering (IJRTE)* 1.6, S. 81–86.
- Hilgefort, I. (Sep. 2019). *Data Analyzer – Ad Hoc Query Analysis for BW Queries with SAP Analytics Cloud*. <https://blogs.sap.com/2019/09/04/data-analyzer-ad-hoc-query-analysis-for-bw-queries-with-sap-analytics-cloud/>, (besucht am: 15.01.2022).

- Holthaus, C., Y.-k. Park und R. Stock-Homburg (Sep. 2015). "People Analytics und Datenschutz—Ein Widerspruch?" In: *Datenschutz und Datensicherheit - DuD* 10, S. 676–681.
- Hölzing, J. A. (2008). *Die Kano-Theorie der Kundenzufriedenheitsmessung. Eine theoretische und empirische Überprüfung*. 1. Aufl. Gabler Verlag.
- Hostmann, B., N. Rayner und T. Friedmann (Okt. 2006). "Gartner's Business Intelligence and Performance Management Framework". In: *Gartner Research*.
- International Organization for Standardization (Dezember 2018). *ISO 30414 Human resource management — Guidelines for internal and external human capital reporting*. ISO Standard. Genf, Schweiz.
- Isson, J. P. und J. S. Harriott (2016). *People Analytics in the Era of Big Data. Changing the Way You Attract, Acquire, Develop, and Retain Talent*. 1. Aufl. Wiley-VCH Verlag.
- Iveta, G. (März 2012). "Human Resources Key Performance Indicators". In: *Journal of Competitiveness* 4.1, S. 117–128.
- Jesse, H. (Jan. 2020). *Kano-Modell: So erzielst du den Wow-Effekt bei deinen Kunden*. <https://dmexco.com/de/stories/kano-modell/>, (besucht am: 18.01.2022). Koelnmesse GmbH.
- Kavis, M. J. (2014). *Architecting the cloud. Designing decisions for cloud computing service models*. 1. Aufl. Wiley-VCH Verlag, S. 9.
- Kemper, H.-G., H. Baars und W. Mehanna (2010). *Business Intelligence – Grundlagen und praktische Anwendungen. Eine Einführung in die IT-basierte Managementunterstützung*. 3. Aufl. Vieweg+Teubner Verlag, S. 4.
- Lindemann, U. (2009). *Methodische Entwicklung technischer Produkte. Methoden flexibel und situationsgerecht anwenden*. 3. Aufl. Springer Verlag.
- Liu, F. u. a. (Juli 2010). "2010 IEEE 3rd International Conference on Cloud Computing". In: *2010 IEEE 3rd International Conference on Cloud Computing* (Miami, Florida). IEEE, S. 402–409.
- Loi, M. (Apr. 2021). *People Analytics muss den Menschen zugutekommen. Eine ethische Analyse datengesteuerter algorithmischer Systeme im Personalmanagement*. https://www.boeckler.de/fpdf/HBS-007994/p_study_hbs_450.pdf. (besucht am 30.01.2022).
- Loscher, G. (2021). *Quick Guide People Analytics. Wie Sie das Personalmanagement verändern können*. 1. Aufl. Springer Verlag.
- Microsoft Corporation (Jan. 2022). *Power BI - Was ist ein lokales Datengateway?* <https://docs.microsoft.com/de-de/power-bi/connect-data/service-gateway-onprem>, (besucht am: 17.01.2022).
- Mircea, M., B. Ghilic-Micu und M. Stoica (Jan. 2011). *Combining Business Intelligence with Cloud Computing to Delivery Agility in Actual Economy*. https://www.researchgate.net/publication/290547197_Combining_business_intelligence_with_cloud_computing_to_delivery_agility_in_actual_economy. (besucht am 30.01.2022).
- Muntean, M. (2015). "Considerations Regarding Business Intelligence in Cloud Context". In: *Informatica Economică* 19.4, S. 55–67.
- National Institute of Standards and Technology's (2011). *Final Version of NIST Cloud Computing Definition Published*. <https://www.nist.gov/news-events/news/2011/10/final-version-nist-cloud-computing-definition-published> (besucht am: 15.01.2022).
- Olszak, C. M. (Jan. 2014). "Business Intelligence in Cloud". In: *Polish Journal of Management Studies* 10.2, S. 115–125.
- Otto, B. und H. Österle (2016). *Corporate Data Quality. Voraussetzung erfolgreicher Geschäftsmodelle*. 1. Aufl. Springer Verlag, S. 18–20.
- Peeters, T., J. Paauwe und K. Van De Voorde (2020). "People analytics effectiveness: developing a framework". In: *Journal of Organizational Effectiveness: People and Performance* 7.2, S. 203–219.
- Pentzek, R. und A. Espig (Juni 2016). "HR and analytics: why HR is set to fail the big data challenge". In: *Controlling & Management Review* 30.3, S. 18–23.
- Reindl, C. und S. Krügl (2017). *People Analytics in der Praxis. Mit Datenanalyse zu besseren Entscheidungen im Personalmanagement*. 1. Aufl. Haufe-Lexware GmbH Co. KG.
- Reindl, C. U. (Juni 2016). *People Analytics: Datengetriebene Mitarbeiterführung als Chance für die Organisationspsychologie*. <https://link.springer.com/article/10.1007/s11612-016-0325-7>. (besucht am 30.01.2022).
- Roth, M. und B. O. Bernardy (Nov. 2019). "Vorhersage von Anforderungen durch Verknüpfung von Szenario-Technik und Kano-Modell". In: *15. Symposium für Vorausschau und Technologieplanung* (Berlin). Hrsg. von J. Gausemeier, W. Bauer und R. Dumitrescu. Heinz Nixdorf Institut, S. 295–310.
- SAP SE (2020a). *SAP Analytics Cloud Help - Enable a Custom SAML Identity Provider*. <https://help.sap.com/viewer/00f68c2e08b941f081002fd3691d86a7/release/en-US/3651184dad944aa2b361ad029a7a8cae.html>. (besucht am: 15.01.2022).
- (Juni 2020b). *SAP Analytics Cloud, analytics designer Developer Handbook*. https://d.dam.sap.com/a/3Y16uka/DeveloperHandbookSACAnalyticsDesigner_v12.0_final.pdf. (besucht am 30.01.2022).
- Satyanarayana, S. (Dezember 2012). "Cloud Computing : SaaS". In: *GESJ: Computer Science and Telecommunications* 36.4, S. 76–79.
- Sauerwein, E. (2000). *Das Kano-Modell der Kundenzufriedenheit. Reliabilität und Validität einer Methode zur Klassifizierung von Produkteigenschaften*. 1. Aufl. Springer Fachmedien Wiesbaden GmbH.
- Savu, L. (Mai 2011). "Cloud Computing. Deployment models, delivery models, risks and research challenges". In: *2011 International Conference on Computer and Management* (Wuhan, China). IEEE, S. 1–5.

- Sharda, R., D. Delen und E. Turban (2015). *Business Intelligence and Analytics: Systems for Decision Support*. 10. Aufl. Pearson Education Inc., S. 11.
- Statista Technology Market Outlook (Oktober 2021). *Enterprise Software Report 2021*. <https://de.statista.com/statistik/studie/id/102693/dokument/software-report/>. (besucht am 30.01.2022).
- Sternad, D. und G. Mödritscher (2018). *Qualitatives Wachstum. Der Weg zu nachhaltigem Unternehmenserfolg*. 1. Aufl. Gabler Verlag, S. 112–114.
- Wille, R. (Juli 2016). “Potenziale und Herausforderungen von People Analytics Einsatzmöglichkeiten für Industrieunternehmen in der DACH-Region”. Bachelorarbeit. Fachhochschule Vorarlberg.
- Zeng, L. u. a. (Okt. 2006). “Techniques, Process, and Enterprise Solutions of Business Intelligence”. In: *2006 IEEE International Conference on Systems, Man, and Cybernetics* (Taipei, Taiwan). IEEE, S. 4722–4726.

Human-Machine Collaboration in Decision-Making

Alexander Hatz

Pforzheim University
Tiefenbronner Straße 65
75175 Pforzheim
hatalex@hs-pforzheim.de

Jurij Jaklič

University of Ljubljana
Kardeljeva ploščad 17
1000 Ljubljana
jurij.jaklic@ef.uni-lj.si

Frank Morelli

Pforzheim University
Tiefenbronner Straße 65
75175 Pforzheim
frank.morelli@hs-pforzheim.de

KEYWORDS

Machine Learning, Human-Machine Collaboration, Collaborative Intelligence, Success Factors

ABSTRACT

The use of ML-based decision support systems in business-related decision-making processes is a proven approach for companies to increase process performance and quality. To a certain extent, machines are capable of reproducing the cognitive intelligence of humans in specific domains. In order to leverage the resulting potential, synergistic human-machine collaboration (HMC) is becoming increasingly important for companies. However, orchestrating HMC is dependent on a set of framework conditions that determine the success of the collaboration. This study¹ examines the research question of how to utilize the concept of collaborative intelligence (CI) to enhance decision-making processes while using machine learning (ML) -based data prediction. The purpose is to identify success factors in the development, design, and implementation of an ML-based predictive analytics solution to orchestrate HMC in decision-making processes. These success factors state recommendations for companies to fulfil the necessary framework conditions for synergistic HMC orchestration. In total, five success factors were identified that represent a combination of theoretical findings and empirical insights. At the same time, further research needs were uncovered, which point out starting points for future research projects in the field of HMC.

INTRODUCTION

The business environment is changing rapidly, especially with the emergence of innovative technologies as an important influencing factor (Vial, 2019, pp. 118-119). The effect of such technologies is two-folded. On one hand, they embody disruptive forces that challenge companies to expand their existing IT infrastructure and to invest in new information systems to avoid being outpaced by competitors in the long run (Li, Porter, Suominen, 2017, pp. 286-287). On the other hand, they are key for developing new, digital strategies to cope with digital transfor-

mation and the challenges it poses. To meet these challenges, automated and scalable business processes must be enriched with data insights and smart technologies (Vial, 2019, p. 122). In this way, the disruptive forces of technology are able to be transformed into potential. Companies that master the digital transformation can emerge as digital pioneers with innovative business processes and permanently change the existing competition on the market in their favor (Jafari-Sadeghi et al., 2021, p. 108).

Machine learning, as a sub-area of artificial intelligence (AI), is an emerging technology that has already disrupted the business environment in the past and will continue to do so in the future (Statista, 2020; Chang, 2020, pp. 99, 110). Due to its versatility and scalability, ML opens up potential in a wide scope of application (Chen & Guestrin, 2016, p. 1). Among other areas, it is used in medicine to diagnose breast cancer, in cybersecurity to detect malware, or in the industry as the basis for business process improvements (Wilson & Daugherty, 2018, pp. 116, 121). Especially in business-related decision-making processes, ML can be used for predictive analytics to forecast outcomes (Nyce, 2007, p. 1). This improves both the flexibility and quality of operational and strategic business decisions to gain a competitive advantage as a company (Kaparthi & Bumblauskas, 2019, p. 660). Since ML solutions are capable of reproducing the cognitive intelligence of humans for abilities related to systematic thinking and pattern recognition, some companies assume that humans are no longer needed in these application areas (Jarrahi, 2018, pp. 577-578; Metcalf, Askay, Rosenberg, 2019, pp. 102-103). A replacement by machines is supposed to be a cost-efficient alternative that achieves the same or even improved performance. However, this approach only leads to short-term success (Kaplan & Haenlein, 2018, p. 24; Malone, 2018; Wilson & Daugherty, 2018, p. 126). The assumption of companies that ML applications are plug-and-play solutions that are implemented once and can solve tasks independently, as humans did before, is a misconception. The data basis of ML-based solutions must be continuously improved, extended, and controlled by humans (Fountaine, McCarthy & Saleh, 2019, p. 64; Ridhawi et al. 2018, pp. 375-376). Furthermore, ML solutions are

¹ The study was conducted in collaboration with SAP SE, a German software development company, and the universities of Ljubljana and Pforzheim.

applications that have been trained for a specific application domain. Outside this scope, the cognitive abilities are not sufficient for adequate problem solving, whereas humans are able to intuitively adapt to changing conditions and find innovative solutions (Kaplan & Haenlein, 2018, p. 16; Malone, 2019, p. 126). Consequently, human intervention is required to maintain the machines performance, which in turn suggests that simply replacing humans with machines is not an appropriate strategy in the long-term. Rather, HMC must be orchestrated, otherwise the desired increase in performance will evaporate (Ridhawi et al., 2018, pp. 375-376; Wilson & Daugherty, 2018). In the field of the CI knowledge base, this problem is addressed and HMC is presented as an alternative outcome. In this way, companies can use the individual strengths of both agents to compensate for the individual weaknesses (Epstein, 2015, pp. 40-41; Wilson & Daugherty, 2018, p. 116).

This paper addresses the research question of how to utilize the concept of CI to enhance decision-making processes while using ML-based data prediction. The purpose is to identify success factors in the development, design, and implementation of an ML-based predictive analytics solution to successfully orchestrate HMC in decision-making processes. The research object is narrowed down to analytical AI systems that are used for the cognitive augmentation of human capabilities. Conversely, this means that industrial robots used for physical augmentation are outside the scope of this work because these types of intelligent machines are usually used separately from humans, which means that no HMC takes place.

As research framework of the paper, a design science research (DSR) approach was used to develop the design artifact. As evaluation method, a case study was implemented within the DSR approach. To collect data from the case study, a multiple sources of evidence collection approach was used.

The first chapter of the paper presents the underlying theoretical knowledge base and is divided into the topics of predictive analytics and CI. Then the developed research design is presented by explaining the used methodology in detail. Based on this, the development and evaluation process of the artifact is presented. Chapters five and six present and critically reflect the research results before the final chapter concludes the paper.

KNOWLEDGE FOUNDATION

Predictive analytics is an established concept that represents the foundation for the development of ML-based predictive analytics solutions. In direct comparison, CI is much less known by name but combines characteristics of research areas such as human-in-the-loop (HITL), explainable artificial intelligence (XAI), and collective intelligence. To define the framework that applies to the design and implementation of the ML-based predictive

analytics solution, an interdisciplinary application of both research areas is necessary.

Predictive Analytics

The need to predict outcomes, budgets, demand, or supplies is crucial for surviving in today's highly competitive business environment (Nyce, 2007, p. 1). Data-driven decision-making processes supported by intelligent systems have become an integral part of many industries (Chen & Guestrin, 2016, p. 1). The availability of sufficient data as well as the progress and simplified access to advanced technologies are driving this trend forward (Nyce, 2007, p. 2). As a basis for predictive analytics solutions, AI in the form of ML models enjoys great popularity. The applications are particularly attractive because they are able to process large amounts of data efficiently and recognize complex relationships within the abundance of data. Furthermore, the ML models are scalable and versatile due to their learning ability (Bawack, Wamba & Carillo, 2021, p. 646).

Collaborative Intelligence

Due to the technical progress in the field of AI, a new paradigm within the IS discipline was initiated in the 2000s, focusing on research of cognitive computing systems. In this new paradigm, modern AI systems are characterized by experiential learning rather than repetitively reproducing programmed knowledge. Furthermore, modern AI systems should be able to interact with humans and share tasks as if they were conscious beings themselves (Aleksander, 2004, pp. 24-25; Epstein, 2015, p. 39).

Collaborative intelligence as a school of thought deals with the interaction between humans and machines and considers replacing humans with innovative ML-based solutions to improve business performance as an outdated mindset (Epstein, 2015, p. 40; Wilson & Daugherty, 2018, p. 116). Instead, the integration of such technologies is about orchestrating HMC in a way that the strengths of both agents augment each other. Accordingly, there is no replacement of humans by machines, but a shift in responsibilities and task allocation between both agents. The capabilities gained using ML augment the existing capabilities of the humans, vice versa, humans complement the machines outside its scope of application as controlling agents (Paschen, Wilson & Ferreira, 2020, p. 412; Bawack, Wamba & Carillo, 2021, pp. 646-647; Epstein, 2015, p. 44).

A functioning HMC is subject to certain framework conditions that must be fulfilled in order to be able to use the potential of both agents (Alizadeh et al. 2020, pp. 4-5; Epstein, 2015, p. 40). However, this aspect is often neglected when integrating innovative technologies into existing business processes, so that the AI adoption fails (Wilson & Daugherty, 2018, p. 116). The literature on CI as well as related research areas collective intelligence,

XAI, and HITL address this challenge and describe recommendations for action in order to create the necessary framework conditions and carry out HMC orchestration (Arrieta et al. 2020, p. 100; Metcalf et al., 2019, pp. 84-86; Wilkens, 2020, pp. 258, 261-262). In the following, these recommendations are summarized and presented as five Design Principles. In the literature, the term Design Principle is not used in the context of CI. Rather, it is an umbrella term introduced by the authors of this paper in order to be able to cluster individual recommendations for action according to their focus area.

Accuracy for the Applicability of the model – Adequate accuracy of a prediction is a basic prerequisite for the solution to add any value at all in practice (Arrieta et al., 2020, p. 100; Shearer, 2000, p. 15). However, accuracy alone is not sufficient to develop an applicable decision support system (Lossos, Geschwill & Morelli, 2021, pp. 316-317). The trade-off between accuracy and transparency is crucial when it comes to optimizing HMC orchestration (Arrieta et al., 2020, p. 100).

Transparency to achieve acceptance, trust, and interpretability – Transparency comprises the explanatory approaches that are expected to make the decision-making logic and results of an AI solution comprehensible to the user. Consequently, transparency is necessary to counteract the black-box character of AI (Arrieta et al., 2020, p. 83-84; Lossos et al., 2021, p. 305). Especially in the context of decision-making processes, the level of transparency is a decisive criterion on whether an ML solution is used in practice or not (Bohanec, Robnik-Šikonja & Borštnar, 2017, p. 1390). Moreover, the results of Arnold et al. (2006, p. 95) have shown that decision-makers tend to use the recommendations of decision support systems if, in addition to the accuracy of the prediction a cognitive alignment is established. Cognitive alignment is defined as a fit between the decision maker's understanding of the underlying business problem and the decision-making logic of the ML solution (Arnold et al., 2006, p. 94). Consequently, transparency is necessary to establish cognitive alignment, and cognitive alignment is in turn the basic prerequisite for user acceptance and trust in the results of a decision-support system (Epstein, 2015, p. 40, Bohanec et al., 2017, p. 1403). Furthermore, transparency lays the foundation for the interpretability of the results as it provides the explanations for the comprehensibility of the decision-support system (Arrieta et al., 2020, p. 83-84). Gilpin et al. (2019, p. 5) define interpretability as “the ability to explain or to present in understandable terms to a human”. The goal of interpretability in an ML context is to describe the underlying logic of the system in a comprehensible way in order to enable decision-makers to derive the correct business decision (cognitive augmentation) as part of an informed decision-making process (Gilpin et al., 2019, p. 2; Lossos et al., 2021, p. 305). There are various recommendations for action to implement the Design Principles. Among other things, developers should carry out a target-group-oriented de-

sign of the artifact, adapted to the expertise of the decision-maker as the addressee of the decision-support system (Harbers, van den Bosch & Meyer, 2010, p. 132). Furthermore, the results of Bohanec et al. (2017) and Lossos et al. (2021) have shown that providing the decision-maker with insights into the development process and decision-making logic of the ML model is a proven way to increase transparency, trust, acceptance, and interpretability. Analogous to Lossos et al., (2021, pp. 312-313) the insights into the solution-finding process can be divided according to three different points in time: Antehoc with regard to input data, AI design with regard to solution development, and post-hoc with regard to output data. Moreover, according to Gilpin et al. (2019, p. 9), the supply of metrics is a proven means to obtain trust and acceptance with the user. However, the choice of metrics needs to be adapted to the individual use case and the addressees.

Participation to foster organizational learning – Viewing ML solution as a simple tool that takes over the part of the work that humans are unwilling or unable to do is outdated. Instead, machines should be seen as an equal partner that elevates the collaboration to a higher level (Demetis & Lee, 2018, pp. 946-947; Schuetz & Venkatesh, 2020, p. 3). Staging human-machine interaction fosters organizational learning in dealing with ML solutions, which in turn strengthens the performance of both and enables informed decision-making (Ansari, Erol & Sihn, 2018, p. 117; Bohanec et al., 2017, p. 1403). There are different recommendations for action to implement the Design Principle. The basis of a successful collaboration is a dialogue in order to achieve the same understanding and a cognitive alignment of the underlying business problem. Such dialogue can be achieved in HMC by providing information about the underlying business problem (feedforward) and explanations (feedback) that make the ML development and decision-making process more transparent to the user (Arnold et al., 2006, pp. 81, 95; Epstein, 2015, p. 44). Furthermore, Chander et al. (2018) and Langley et al. (2017) describe the backward integration of business users into the ML development process as a way to intensify the interaction between humans and machines. The idea is based on giving the user access to the development process, thereby considering their perspective on the underlying business problem at an early stage to achieve cognitive alignment (Arrieta et al., 2020, p. 87; Chander et al., 2018, p. 4).

Task allocation due to competence transferability – A successful group performance is based on a reasonable allocation of responsibilities based on individual competencies. Thus, the cognitive augmentation of humans by machines is also dependent on the premise that the shared task is divided into sub-tasks, and these are allocated based on the individual competence profiles (Ansari et al. 2018, p. 119). Through technological progress in the field of AI, a transfer of competencies between humans and machines is taking place. AI is able to adapt the cognitive abilities of humans, which promotes the integration of

machines into formerly human-centric processes. Humans, in turn, must develop new competencies in an AI environment. This shift requires the adaptation of individual competence profiles and a clear definition of responsibilities based on revised competence profiles (Ansari et al., 2018, pp. 121-122; Wilkens, 2020, pp. 260-262). Based on the competence profiles of humans and machines, specific task pools (human-specific, machine-specific, shared task) should be defined individually adapted to the use case at hand and employed as a basis for task allocation (Ansari et al. 2018, p. 119; Wilkens, 2020, p. 257).

Governance to ensure a regulatory framework – Governance represents the overarching regulatory framework to which all Design Principles should be subject (Lossos et al., 2021, pp. 307-308). In ML-based decision support systems, data security and privacy must be ensured, legal frameworks have to be respected and discrimination against individuals must be avoided (Arrieta et al., 2020, p. 84, European Commission, 2019). Similarly, the implementation of control mechanisms with human oversight is a way to define responsibilities and establish the regulatory framework. Analogous to the HITL concept, interactive human involvement ensures the objectives (European Commission, 2019, pp. 19-20). Furthermore, for in-depth insights, the following frameworks can be recommended in the context of IT governance: COBIT 2019 (ISACA, 2019), the General Data Protection Regulation (GDPR) (Sartor, 2020), or "Assessment List for Trustworthy AI" (European Commission, 2019).

RESEARCH DESIGN

The research design consists of a combination of the DSR methodology and a case study, whereby the case study functions as evaluation method within the DSR framework. By describing the scientific and empirical techniques used, the validity and reliability of the results presented can be questioned critically.

Design Science Research

The DSR framework of Hevner et al. (2004), presented in Figure 1, was used as the overarching research methodology of the study. Design science is an iterative development and evaluation process of a design artifact synthesized from concepts and methodologies of an existing knowledge base (rigor) to solve existing business needs (relevance) of an organization (Hevner et al., 2004, pp. 76-78). The underlying knowledge base refers to the research areas of predictive analytics and CI as well as the related topics of XAI, HITL, and collective intelligence to ensure a holistic perspective. The relevance of the DSR approach is embodied by the business needs of organizations and people facing the challenge of integrating analytic AI artifacts into enterprise infrastructure as part of the digital transformation of decision-making processes (van der Merwe, Gerber, Smuts, 2020, p. 170; Hevner et al. 2004, p. 85).

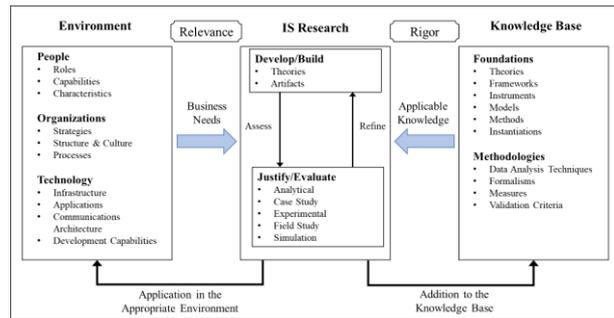


Figure 1: Information Systems Research Framework [Source: Hevner et al., 2004, p. 80.]

As DSR artifact, a prototype of a decision support system in form of a predictive analytics solution was developed, which consists of an ML-based prediction model and a dashboard as an interface for HMC. To implement the iterative design and evaluation loop according to Hevner et al. (2004, p. 85), a case study served as the evaluation method. However, the iterative design and evaluation loop is limited to one cycle. As evidence collection methods of the case study, a mixed method in the form of a triangulation of a survey and semi-structured interviews were conducted. This allows to collect quantitative as well as qualitative insights for the evaluation of the design artifact. To validate the gathered data, a pattern matching approach was carried out. A detailed description of the mixed method evaluation and the pattern matching approach will be given in the presentation of the case study. As a conclusion to the DSR process, the insights gathered in the design and evaluation process serve as research contributions to practice and theory (Baskerville et al., 2018, pp. 367-368; Hevner et al., 2004, pp. 80-81).

Case Study

The case study was created following a critical single case design. According to Yin (2018, p. 46), the use of a clearly defined theoretical foundation is particularly important in single case design to ensure external validity. Therefore, concepts of the research areas presented in the literature review were logically linked to define research propositions. Based on Yin, the research propositions in Table 1 serve to determine the research direction of the research question to ensure that evidence is sought in the right places (2018, pp. 27-28).

Table 1: Research Propositions [source: own work.]

1	CI ensures a focus on the HMC when implementing a predictive analytics solution.
2	The interdisciplinarity of the CI Design Principles ensures a holistic perspective on the basic requirements of an HMC orchestration.
3	The implementation of analytical AI artifacts using the CI Design Principles ensures an informed decision-making process.

The case unit was the fictitious company named Global Bike Incorporated (GBI) which is from the SAP University Alliances. The SAP University Alliance provides a framework of SAP software and learning content used for educational institutes for conducting case studies (SAP, 2021). The GBI represents a medium-sized company operating in the discrete manufacturing sector. In the long run, they are pursuing the business goal of increasing their competitiveness in the market by establishing a bike-sharing business as a second major sales channel. To this end, the management of the GBI aims to develop a predictive analytics solution in the form of a decision support system. The added value of the decision support system is to provide a demand forecast as well as background information and influencing factors of customer behavior for the responsible managers in order to improve the performance of operational and strategic decisions. By implementing CI Design Principles, the specific design of the decision support system is intended to orchestrate HMC to add value to the solution. The data used for the case study originates from a data set of Kaggle, which is based on a real bike-sharing use case (Kaggle, 2022).

To collect data from the case study, a multiple sources of evidence collection approach was used to increase the construct validity and reliability of the results (Yin, 2018, p. 126). For this purpose, the design artifact was tested in a simulation in the form of a usability test (Hevner et al., 2004, p. 80). The results of the usability test were collected through a combination of survey and semi-structured interviews. The triangulation ensured the validity, reliability, and practicality of the collected data (Yin, 2018, pp. 118, 126-128). As guidance, Saunders et al. (2016) and Yin (2018) were used for a scientifically correct methodology. The collected data was analyzed using a pattern matching approach. A theoretical pattern is a hypothesis about what is expected in the observational realm – observed patterns. The pattern matching approach ensures a structured research process that increases the internal validity of the results (Sinkovics, 2018, p. 2; Yin, 2018, pp. 175-178). Based on the findings of Sinkovics (2018), it is up to the researcher to decide in what form and granularity the theoretical patterns for pattern matching are defined. Therefore, the theoretical patterns in Table 2 were derived from the research propositions to establish an alignment between the theoretical knowledge base, research question, and the case study. As part of the derivation process, the scope of the theoretical patterns was delimited to AI in the form of ML and decision processes. This ensures that the theoretical patterns and the observed patterns have the same level of abstraction.

Table 2: Theoretical Patterns [source: own work.]

1	The applicability of a decision-making system is linked not only to accuracy but also to the transparency of the underlying ML model.
---	---

2	Achieving cognitive alignment between humans and machines is a precondition for establishing trust and acceptance within the HMC.
3	An interpretable and comprehensible design of a decision-making system determines the added value of cognitive augmentation.
4	The implementation of measures that orchestrate the interaction between humans and machines promotes mutual learning and increases overall performance.
5	The implementation of ML-based solutions in decision-making processes requires an adjustment of existing roles and responsibilities based on the capabilities of the participants.
6	Governance as an overall regulatory framework in dealing with decision-making systems is needed to ensure compliance requirements in HMC.

The data collected from the survey and semi-structured interviews serve as basis for the empirically observed patterns. Together, the theoretical and observed patterns form the counterparts for the pattern-matching approach, which enables the theorizing process of the success factors. A sample of three people, two business users, and one expert user were selected for the data collection. The usability testing with the survey and the semi-structured interview together took about an hour per participant. The observation units were SAP employees representing potential users of a decision support system. When selecting the participants, it was taken care that they had no direct relationship to the research project to minimize the influence of possible bias.

ARTIFACT DEVELOPMENT

Within the case study, the predictive analytics solution was implemented to simulate the orchestration of the HMC. The aim was to create a synergy between humans and machines through a cognitive augmentation of the decision-maker by the machine, in order to enable an informed decision-making process. In the backend, an ML model formed the basis for calculating the demand forecast of the decision support system. The front end was represented by a dashboard that acts as a user interface for decision support.

For the development of the ML model, the methodology of the CRISP-DM process model was followed as a blueprint. Since this is a de-facto standard, the necessary rigor of the DSR approach was also ensured. In the selection of the ML method, special attention was paid to transparency and interpretability in addition to the accuracy of the model. Therefore, several models were compared with each other in order to achieve an optimal trade-off between transparency and performance.

For the development of the dashboard, the recommendations of Few (2006) were followed to ensure a coherent design, an appropriate information density, and an optimal visualization adapted to the specific content. While selecting the content to be visualized, measures were derived from the Design Principles in addition to decision-relevant information. The measures provide information that contextualizes the demand forecast and thus indirectly contribute to decision support by helping the decision-maker to process the information.

The dashboard is divided into several pages. This corresponds to the basic idea of decomposability as a measure to reduce the complexity of the decision support system. The division into logically related subject areas is intended to increase comprehensibility and transparency for the user. Despite the division, the use of links and drill-down options ensures that the content is logically linked. This creates a story flow that makes it easier for the user to process the content provided. Furthermore, each page contains textual explanations that are directly linked to the corresponding key figures or diagrams. In terms of a target group-oriented design, this fosters transparency and increases comprehensibility for better interpretability for both novice and experienced users. For the same purpose, a help button has been implemented to support the use of the dashboard.

Landing Page - It is the user's first point of contact with the predictive analytics solution. The page contains an overview of all other pages of the dashboard including a short description and a link to the corresponding pages.

Decision Support Page - The decision support page serves as the central information base for decision support in controlling and planning dynamic business processes within the bike-sharing business. In Figure 2 a simplified depiction of the dashboard page is provided.

Introductory Sentence		Link Help Page
Forecast	Feature Importance – Explained Variance	
Performance Metrics	Link Calendar Effects	Link Weather Effects

Figure 2: Decision Support Page

On the left half of the page, the demand forecast is displayed which can be used by the decision-maker as decision support to draw conclusions for future demand planning. To show the user that the forecast results are trustworthy and reliable, the ML model performance metrics are visualized directly below the forecast. This is a model-agnostic measure intended to provide transparency post hoc to the model development process. On the right half of the page, the underlying feature importance of the two ML models is visualized by ranking the top ten features per model that explain the largest share of the data variance in the forecast. This post-hoc model-agnostic measure not only helps the decision-maker to gain transparency about the decision logic of the ML model but also contains additional decision-relevant information for the underlying business problem.

Knowledge Foundation Page - To ensure maximum transparency for the decision-maker, the knowledge foundation page provides basic information about the features as well as insights into the data set on which the ML model development is based (see Figure 3). Those ante-hoc feedforward measures allow novice and expert users to gain a holistic understanding of the data set and contextualize the forecast to ensure informed decision-making.

Introductory Sentence		Link Help Page
Textual Feature Description	Descriptive Analytics Metrics	

Figure 3: Knowledge Foundation Page

For this purpose, the left half of the page defines each feature from the underlying dataset and the right half of the page visualizes a selection of relevant descriptive metrics for the dataset.

In-depth Insights Pages - The two in-depth insights pages provide the user with the opportunity to understand the decision logic of the ML model by showing calendar and weather effects within the dataset that the algorithm uses for prediction (see Figure 4 and 5).

Introductory Sentence		Link Help Page
Number of Customers per Year	Distribution of Customers per Weekday	

Figure 4: Calendar Effects Page

Introductory Sentence		Link Help Page
Correlation of Temperature to the Number of Customers	Correlation of Humidity and Wind Speed to the Number of Customers	

Figure 5: Weather Effects Page

Particularly as a complement to the visualized feature importance on the decision support page, these ante-hoc insights serve as a way to contextualize and improve the interpretability of the feature importance ranking. Furthermore, these insights act as a cognitive augmentation of the decision-maker by creating a holistic understanding of relevant influencing factors for the underlying business problem.

Machine Learning Decision Making Page - The ML decision-making page serves as a detailed explanation of the decision-making process and the decision logic of the ML model (see Figure 6). Both contribute to the creation of transparency of the ML design, which in turn increases trust and acceptance in the ML model. Furthermore, valuable insights into the underlying business problem can be derived from the decision logic of the ML model.

Introductory Sentence		Link Help Page
Feature Importance – based on F-Score	Visualized Regression Tree	

Figure 6: Machine Learning Decision Making Page

As a model-agnostic measure the feature importance based on the F-score is provided on the left half of the page. Again, the visualization form of a ranking of the feature per model is used for this purpose. Compared to the feature importance based on the share of variance explained, the F-score-based feature importance indicates how often a feature is used to split the data across all trees. This should help the user to understand the influence of the metric explanatory variables compared to the categorical variables to understand the decision-making process of the ML model. As a complement to the F-score-based feature importance, the right half of the page includes a complexity-reduced visualization of the regression tree underlying the ML model. The visualization, as a model-specific measure, is intended to help the user make a connection between how the ML model uses the different variables and to generate the results of the forecast.

Evaluation for Continuous Improvement Page - The evaluation for the continuous improvement page serves as a monitoring and data collection mechanism to identify optimization potentials with regard to the orchestration of the HMC when dealing with the predictive analytics solution (see Figure 7).

Introductory Sentence		Link Help Page
Evaluation Metrics	Link to Survey	
	Word Cloud	

Figure 7: Evaluation for Continuous Improvement Page

Therefore, a link to an online survey is provided in the top right-hand corner of the page. The survey acts as a feedback loop in which users can participate to initiate the desired continuous improvement process and promote mutual learning. On the left side of the page, the evaluation metrics are provided. By answering the questionnaire, five different dimensions are quantitatively assessed, which in their entirety are intended to evaluate the performance of the implemented measures and determine the added value of the dashboard. In addition to the quantitative evaluation via metrics, qualitative insights from the survey are visualized via the word cloud on the lower right side of the page.

Help Page - The purpose of the help page is to provide background information on the predictive analytics solution in order to create a transparent knowledge base, which is especially necessary for ML-based solutions to build trust in the solution (see Figure 8).

Introductory Sentence		Link Help Page
Business Problem	Model Design	Governance
Glossary/ Nomenclature		

Figure 8: Help Page

Therefore, information about the underlying business problem is provided to make the framework conditions transparent. In addition, key facts of the model design

and information on the governance of the predictive analytics solution are stated.

ARTIFACT EVALUATION AND DATA ANALYSIS

The two evidence collection methods were carried out separately: As a starting point, each respondent was educated about the underlying business problem for which the DSR artifact was developed within the case study. Afterward, they performed an independent usability test of the decision support systems user interface to gather impressions and experience of using the solution before answering the questions in the survey. Following the survey, each respondent was invited to a semi-structured interview. Complementing the survey with the interviews was providing a deeper understanding of the motivations behind the respondent's answers and allow contextualization results to enrich the quantitative information with qualitative ones.

A descriptive analysis of the survey was conducted at question block and individual question level to draw conclusions for the observed patterns. When analyzing the qualitative data from the semi-structured interviews, a thematic analysis according to Saunders et al. (2016, pp. 579-580) was conducted. This involved looking for patterns within the data that could be used to contextualize the quantitative data from the questionnaire. The same sequence of questions in the questionnaire and the interview implied that there was already a thematic coherence, which simplified the analysis. The findings of the data analysis regarding the subjects' perception of the underlying mechanisms of the CI Design Principles form the basis of the observed patterns. Within the pattern matching approach, these observed patterns were compared with the theoretical patterns.

The pattern matching showed that in most cases there is an agreement between theory and practice. Patterns one, two, three, four, and six could be matched. Pattern five could not be matched because the use case of the case study was without an implementation scenario in a real environment. Due to that fact, no data could be collected for an observed pattern, which does not allow an empirically valid statement about a possible match. Nevertheless, the pattern matching provided valuable insights from theory and practice for the definition of success factors. In addition, future research needs were identified.

DISCUSSION

The present success factor research aimed to identify operational factors that distinguish successful from less successful HMC orchestrations. It builds on the assumption that success or failure can be traced back to central influencing factors that have a decisive impact on HMC orchestrations. Due to the increasing relevance of the integration of AI artifacts in the form of ML-based solutions in decision-making processes and the associated complexity of HMC orchestration, success factors are required to prevent AI adoptions from failing. The success

factors were clustered according to a thematic coherence to the individual patterns of the pattern matching approach. In terms of content, they comprise the quintessence of theoretical and practical insights derived from the data analysis.

Ensure Transparency - To ensure transparency, the solution development process and the decision logic of ML-based solutions must be comprehensible to the decision-maker. A target group-oriented selection of the provided content adapted to the knowledge level of the user is recommended. In general, transparency has a significant influence on all other success factors. Furthermore, each decision-maker needs different components from the solution development process to understand the functionality and decision-making process of the ML-based solution. The selection of the appropriate measures is linked to factors such as the decision maker's individual statistical background knowledge or experience in dealing with ML solutions. At the same time, personal preferences also play a decisive role. Therefore, as a success factor, the target group and its characteristics must be examined in advance of the design of the solution. The division of the components of the ML-based solution into easily digestible building blocks has proven to be a successful measure for reducing complexity. Hereby, it is success-critical that the user can independently link the logical coherence of the individual building blocks by putting them together in a compelling story. When selecting measures, recommended actions include providing the results of descriptive analysis of the data set to provide identified patterns, correlations, or metrics that facilitate a general understanding of the underlying business problem. This makes the solution development process transparent to the decision-maker and allows contextualization of the information. The information needed for this can be generated by following an accepted process model in the ML design process, such as the CRISP-DM model. Besides the visualization of quantitative insights into the decision logic of ML models, the provision of qualitative insights in the form of textual explanations has proven to be useful. The textual explanations are a flexible complement to visualizations or indicators to provide additional information. In particular, enriching model-agnostic and model-specific measures with textual explanations has proven to be best practice in decision support system design.

Foster Trust and Acceptance - The usability of the decision support system depends on whether humans trust the provided information and thus accept the collaboration with the machine. It is in the nature of human beings to question things critically and to check whether they are in line with their understanding and views before they accept them. Therefore, achieving a cognitive alignment between the decision maker's understanding of the underlying business problem and the ML solution finding process is key to success in building trust and acceptance. In this context, ensuring transparency is success critical, be-

cause the results of the case study have shown that cognitive alignment builds upon transparency. A transparent design is a prerequisite for the decision-maker to be able to identify and resolve the cause of a possible discrepancy in cognitive alignment. Consequently, from each phase of the solution development (ante-hoc, ML design, post-hoc), holistic information must be provided in order to build trust and acceptance. In particular, the illustration of correlation and interaction effects in the data set through explorative statistical methods combined with insights into the decision logic of the ML solution through a post-hoc analysis have proven to be effective here. Furthermore, the process of cognitive alignment can be accelerated by integrating the decision-maker backward into the development process or by providing contextual information and insights to the developer from the ML design phase.

Enable Interpretability - The added value of a decision support system depends on whether the decision-maker is able to process the results of the decision-support system as well as the provided explanations in order to incorporate it into the decision-making process. Under these conditions, the decision support system fulfills its purpose of serving the decision-maker as cognitive augmentation in the sense of HMC to conduct an informed-decision making process. The results of the case study have shown that shortcomings in the implementation of the success factors of transparency, as well as trust and acceptance, have a negative impact on interpretability. Transparency provides the necessary knowledge base through the target group-oriented provision of information. Establishing a cognitive alignment is crucial for the cognitive augmentation of the decision support system to be trusted and accepted by the user. For achieving interpretability, the explanations regarding the ML decision-making logic are more relevant for cognitive augmentation and conducting an informed decision-making process than insights into the ML development process. The combination of model-agnostic and model-specific measures, as well as the provision of descriptive statistical analysis in the form of influencing factors, correlations, and metrics, provides the necessary transparency concerning interpretability. Moreover, the integration of textual explanations as a complement to the other measures provides the reliability needed to avoid misinterpretations, to ensure that no information is lost and that the correct conclusions can be drawn.

Implement interaction-promoting Design - The use of mechanisms such as feedback and optimization loops ensure that the decision-makers actively engage with decision support systems. This triggers mutual learning processes that improve HMC throughout the organization. The results of the case study have shown that the basis for implementing an interaction-promoting design is the provision of background information about the development process and the functioning of the decision support system. This information enables the decision-maker to actively engage with the solution, expand their individual

knowledge level, and identify optimization potential in dealing with the HMC. At the same time, the expanded knowledge also promotes an informed decision-making process, which has a positive effect on the performance of the decision-making process. When implementing a feedback/ optimization loop, it is crucial for success that the underlying communication process between the decision-maker and the developer is transparent, fast, and simple. At the same time, the result must be evaluated and made available transparently so that measures for improvement can be derived.

Providing Governance – A successful HMC orchestration requires an overarching regulatory framework that defines the conditions for collaboration. The case study has shown that the underlying compliance requirements such as data security, data origin, or fairness must be clearly defined and communicated to the decision-maker in an understandable way. Moreover, rules that specify limitations and responsibilities in dealing with decision support systems must be made transparent. Each decision support system has certain limitations that affect the cognitive augmentation of the decision-maker. For example, there are influencing factors that cannot be quantified but must still be considered in the decision-making process. These restrictions need to be pointed out since they have a decisive influence on the quality of the decision support and require the active intervention of the decision-maker. Furthermore, the limitations are included in the definition of areas of responsibility between humans and machines. As another crucial aspect of the regulatory framework, an allocation of responsibilities is necessary to create the basis for HMC. This ensures that subtasks are divided between humans and machines based on their individual strengths.

CRITICAL REFLECTION AND FUTURE RESEARCH

In order to be able to check the applicability of the success factors presented for theory and practice, a critical examination of the external validity of the results must be carried out. For this purpose, the methodological approach is analyzed concerning the collected results. The use of the DSR approach as an overarching research framework has proven to be useful in guaranteeing a scientifically and methodologically recognized approach, which has a positive effect on the external validity of the success factors. Nevertheless, it should be noted that in the sense of the iterative DSR approach, further development and evaluation cycles should follow. In this way, the representativeness of the results can be increased in the long term. The case study as part of the DSR approach has proven to be a useful evaluation method for the present research purpose and research question to generate the necessary data. The single-case design of the case study also achieved the desired effect of being able to test the developed artifact based on the fictitious use case and to demonstrate the relevance of the success factors in the first instance. Only for pattern five, the use case was not

suitable. Therefore, case studies with real implementation scenarios should be investigated in future research projects in order to examine the functioning of the measures associated with pattern five. At the same time, further case studies provide the basis for further evaluation cycles of the DSR approach, which would allow the reproducibility of the success factors and the associated mechanisms to be investigated. In this way, the external validity can be further increased. Concerning the evidence collection methods used, the triangulation of survey and interview proved to be ideal for the present research objective. By combining quantitative findings as a snapshot of the usability testing and qualitative in-depth findings on the motivations of the respondents, a comprehensive database was created that could be used for pattern matching as a method of analysis. The pattern matching approach also fulfilled the desired purpose and was able to establish a link between theory and practice in the context of researching critical success factors. Nevertheless, it must be critically noted that the small number of participants must be taken into account, especially with regard to the validity of the quantitative data. Although the findings per subject were rich and varied, there is still potential for improvement to increase the representativeness of the results in the long term. All in all, it can be assumed that the developed research methodology produced valid, reliable, and representative results to answer the research question. However, the research is not yet completed at this point and should be iteratively advanced based on further research projects, analogous to the DSR approach. In particular, involving additional participants in the evidence collection process can improve the validity of the quantitative database, which in turn increases the representativeness of the identified success factors.

CONCLUSION

Companies that want to use the potential of emerging technologies such as AI in the form of ML solutions to their advantage as part of the digital transformation have to deal with AI adoption. More and more situations will occur in which people are compared with machines. Therefore, companies must ask themselves how they want to implement AI in their company. Companies that use these innovations to replace humans with machines will probably experience only a brief performance boost. Instead, companies that manage to achieve synergetic collaboration between humans and machines will gain a long-term competitive advantage by augmenting the strengths of both agents to compensate for their individual weaknesses.

To provide guidance for these companies, the success factors presented in the paper represent heuristic recommendations for action that can be used to create the necessary framework conditions for successful HMC orchestration. The success factors are a combination of theoretical propositions that were confirmed by the usability testing of the developed decision support system and additionally enriched with empirical findings from the case

study. The addressees are primarily practitioners who are considering the use of ML solutions in the context of digital transformation and need to orchestrate the HMC. Companies that succeed in achieving synergistic collaboration between humans and machines will gain a long-term competitive advantage by amplifying the strengths of humans and machines to compensate for their individual weaknesses.

Based on the current popularity and number of publications in the field of AI, it can be assumed that the relevance and impact for the business environment will continue to increase. Consequently, within the realm of companies attempting to become an intelligent enterprise, the relevance of a successful HMC orchestration will continuously be reinforced. It remains to be seen how HMC will change when current concepts such as Artificial Super Intelligence are implemented in reality and what implications this will have for the collaboration between humans and machines.

LITERATURE

- Alizadeh, F., Esau, M., Stevens, G. & Cassens, L. (2020) eXplainable AI: Take one Step Back, Move Two Steps forward. Workshop on User-Centered Artificial Intelligence (UCAI '20). Retrieved January 30, 2022 from <https://dl.gi.de/bitstream/handle/20.500.12116/33513/muc2020-ws-369.pdf?sequence=1&isAllowed=y>.
- Aleksander, I. (2004). Advances in Intelligent Information Technology: Re-Branding or Progress Towards Conscious Machines? *Journal of Information Technology*, 19(1), 21-27. Doi.org/10.1057/palgrave.jit.2000001.
- Ansari, F., Erol, S. & Sihni, W. (2018). Rethinking Human-Machine Learning in Industry 4.0: How Does the Paradigm Shift Treat the Role of Human Learning? *Procedia Manufacturing*, 23, 117-122. Doi.org/10.1016/j.promfg.2018.04.003.
- Arnold, V., Clark, N., Collier, P. A., Leech, S. A. & Sutton, S. G. (2006). The Differential Dse and Effect of Knowledge-Based System Explanations in Novice and Expert Judgment Decisions. *MIS Quarterly*, 30(1), 79-97. Doi.org/10.2307/25148718.
- Arrieta et al. (2020). Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges toward Responsible AI. *Information Fusion*, 58, 82-115. Doi.org/10.1016/j.inffus.2019.12.012.
- Baskerville, R., Baiyere, A., Gregor, S., Hevner, A., & Rossi, M. (2018). Design Science Research Contributions: Finding a Balance between Artifact and Theory. *Journal of the Association for Information Systems*, 19(5), 358-376. Doi:10.17705/1jais.00495.
- Bawack, R., Wamba, S. & Carillo, L. (2021). A framework for understanding artificial intelligence research: insights from practice. *Journal of Enterprise Information Management*, 34(2), 645-678. Doi:10.1108/JEIM-07-2020-0284.
- Bohanec, M., Robnik-Šikonja, M. & Borštnar, M. (2017). Decision-making framework with double-loop learning through interpretable black-box machine learning models. *Industrial Management & Data Systems*, 117(7), 1396-1406. Doi:10.1108/IMDS-09-2016-0409.
- Chander, A., Srinivasan, R., Chelian, S., Wang, J. & Uchino, K. (2018). Working with beliefs: AI transparency in the enterprise. Workshops of the ACM Conference on Intelligent User Interfaces. Retrieved January 30, 2022 from <http://ceur-ws.org/Vol-2068/exss14.pdf>.
- Chang, S-H. (2020). Technical trends of artificial intelligence in standard- essential patents. *Data Technologies and Applications*, 55(1), 97-117. Doi.org/10.1108/DTA-10-2019-0178.
- Chen, T. & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785–794. Doi: 10.1145/2939672.2939785.
- Demetis, D. S. & Lee, A. S. (2018). When Humans Using the IT Artifact Becomes IT Using the Human Artifact. *Journal of the Association for Information Systems*, 19(10), 929-952. Doi: 10.17705/1jais.00513.
- Epstein, S. (2015). Wanted: Collaborative intelligence. *Artificial Intelligence*, 221, 36-45. Doi.org/10.1016/j.artint.2014.12.006.
- European Commission (2019). Ethics guidelines for trustworthy AI. Retrieved January 30, 2022 from <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>.
- Few, S. (2006). *Information Dashboard Design – The Effective Visual Communication of Data*. Italy. O'Reilly & Associate.
- Fountaine, T., McCarthy, B & Saleh, T. (2019). Building the Ai-Powered Organization. *Harvard Business Review*, 62-73.
- Gilpin et al. (2019). Explaining Explanations: An Overview of Interpretability of Machine Learning. 5th IEEE International Conference on Data Science and Advanced Analytics (DSAA 2018). Turin, Italy: IEEE. Retrieved January 30, 2022 from <https://arxiv.org/pdf/1806.00069v3.pdf>.
- Harbers, M., van den Bosch, K. & Meyer, J.-J. (2010). Design and evaluation of explainable BDI agents. *IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*, 125–132. Doi:10.1109/WI-IAT.2010.115.
- Hevner, A., R., March, S., T., Park, J. & Sudha, R. (2004). Design Science in Information Systems Research. *MIS Quarterly*, 28(1), 75-100. Doi.org/10.2307/25148625
- ISACA (2019). COBIT 2019 Framework: Governance and Management Objectives. Retrieved January 30, 2022 from https://www.isaca.org/bookstore/bookstore-cobit_19-digital/wcb19fgm.
- Jafari-Sadeghi, V. ,Garcia-Perez, A., Candeló, E. & Couturier, J. (2021). Exploring the impact of digital transformation on technology entrepreneurship and technological market expansion: The role of technology readiness, exploration and exploitation. *Journal of Business Research*, 124, 100-111. Doi.org/10.1016/j.jbusres.2020.11.020.
- Jarrahi, M. (2018). Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making. *Business Horizons*, 61, 577-586. Doi.org/10.1016/j.bushor.2018.03.007.
- Kaggle (2022). Bike-Sharing in Washington D.C. Dataset. Retrieved January 30, 2022 from

- <https://www.kaggle.com/marklvl/bike-sharing-dataset?select=Readme.txt>.
- Kaparthi, S. & Bumbalaskas, D. (2020). Designing predictive maintenance systems using decision tree-based machine learning techniques. *International Journal of Quality & Reliability Management*, 37(4), 659-696. Doi.org/10.1108/IJQRM-04-2019-0131.
- Kaplan, A. & Haenlein, M. (2019). Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Business Horizons*, 62, 15-25. Doi.org/10.1016/j.bushor.2018.08.004.
- Langley, P., Meadows, B., Sridharan, M. & Choi, D. (2017). Explainable agency for intelligent autonomous systems. *AAAI Conference on Artificial Intelligence*, 4762–4763. Retrieved January 30, 2022 from <http://www.isle.org/~langley/papers/agency.iaai17.pdf>.
- Li, M., Porter, A., & Suominen, A. (2017). Insights into relationships between disruptive technology/innovation and emerging technology: A bibliometric perspective. *Technological Forecasting & Social Change*, 129, 285-296. Doi.org/10.1016/j.techfore.2017.09.032.
- Lossos C., Geschwill, S. & Morelli, F. (2021). Offenheit durch XAI bei ML-unterstützten Entscheidungen. Ein Baustein zur Optimierung von Entscheidungen im Unternehmen? *HMD*, 58, 303-320. Doi.org/10.1365/s40702-021-00707-1.
- Malone, T. (2018). How Human-Computer 'Superminds' Are Redefining the Future of Work. Retrieved January 30, 2022 from <https://sloanreview.mit.edu/article/how-human-computer-superminds-are-redefining-the-future-of-work/>.
- Malone, T. (2019, January). Published interview conducted by Guszca, J. & Schwartz J. Superminds: How humans and machines can work together. *Deloitte Review*, 24, 121-131. Retrieved January 30, 2022 from <https://www2.deloitte.com/us/en/insights/focus/technology-and-the-future-of-work/human-and-machine-collaboration.html>.
- Metcalf, L., Askay, D. A. & Rosenberg, L. B. (2019). Keeping Humans in the Loop: Pooling Knowledge through Artificial Swarm Intelligence to Improve Business Decision Making. *California Management Review*, 61(4), 84-109. Doi.org/10.1177/0008125619862256.
- Nyce, C. (2007). PREDICTIVE ANALYTICS WHITEPAPER. American Institute for CPU/Insurance Institute of America. Retrieved January 30, 2022 from <http://www.the-digital-insurer.com/wp-content/uploads/2013/12/78-Predictive-Modeling-White-Paper.pdf>.
- Paschen, J., Wilson, M. & Ferreira, J., (2020). Collaborative intelligence: How human and artificial intelligence create value along the B2B sales funnel. *Business Horizons*, 63, 403-414. Doi.org/10.1016/j.bushor.2020.01.003.
- Ridhawi, I., Otoum, S., Aloqaily, M. & Boukerche, A. (2018). Generalizing AI: Challenges and Opportunities for Plug and Play AI Solutions. *IEEE Network*, 35(1), 372-379. Doi.org/10.1109/MNET.011.2000371.
- SAP. (2021). SAP University Alliances. Retrieved January 30, 2022 from <https://www.sap.com/about/company/innovation/next-gen-innovation-platform/university-alliances.html>.
- Sartor, G. (2020) The impact of the General Data Protection Regulation (GDPR) on artificial intelligence. Retrieved January 30, 2022 from [https://www.europarl.europa.eu/RegData/etudes/STUD/2020/641530/EPRS_STU\(2020\)641530_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2020/641530/EPRS_STU(2020)641530_EN.pdf).
- Saunders, M., Lewis, P., & Thornhill, A. (2016). *Research Methods for Business Students* (7th ed.). Harlow, England: Pearson Education Limited.
- Schuetz, S. & Venkatesh, V. (2020). Research perspectives: the rise of human machines: how cognitive computing systems challenge assumptions of user-system interaction. *Journal of the Association for Information Systems*, 21(2), 1-43. Retrieved January 30, 2022 from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3680306.
- Shearer, C. (2000). The CRISP-DM Model: The New Blueprint for Data Mining. *Journal of Data Warehousing*, 5 (4), 13-22. Retrieved January 30, 2022 from <https://mineracaodados.files.wordpress.com/2012/04/the-crisp-dm-model-the-new-blueprint-for-data-mining-shearer-colin.pdf>.
- Sinkovics, N. (2018). Pattern matching in qualitative analysis. In Catherine Cassell, Ann Cunliffe, & G. Grandy (Eds.), *The SAGE Handbook of Qualitative Business and Management Research Methods* (2, 468-485). Thousand Oaks, US: Sage Publications Ltd.
- Statista (2020). Adoption rate of emerging technologies in organizations worldwide as of 2020, by scale. Retrieved January 30, 2022 from <https://www.statista.com/statistics/661164/worldwide-cio-survey-operational-priorities/>.
- Van der Merwe A., Gerber A., Smuts H. (2020) Guidelines for Conducting Design Science Research in Information Systems. In: Tait B., Kroeze J., Gruner S. (eds) *ICT Education*. SACLA 2019. Communications in Computer and Information Science, 1136. Springer, Cham. Doi.org/10.1007/978-3-030-35629-3_11
- Vial, G. (2019). Understanding digital transformation: A review and a research agenda. *The Journal of Strategic Information Systems*, 28(2), 118-144. Doi.org/10.1016/j.jsis.2019.01.003.
- Wilkens, U. (2020). Artificial intelligence in the workplace – A double-edged sword. *The International Journal of Information and Learning Technology*, 37(5), 253-265. Doi.org/10.1108/IJILT-02-2020-0022.
- Wilson, H. J. & Daugherty, P. R. (2018). Collaborative intelligence: Humans and AI are joining forces. *Harvard Business Review*. 96(4), 114-123.
- Yin, R. K. (2018). *Case Study Research and Applications: Design and Methods* (6th ed.). Thousand Oaks, California: SAGE Publications.

Konzeption und Aufbau einer Trainingsinfrastruktur für virtuelle Sensoren

Malte Hoffmann

Technische Informatik

Hochschule
Albstadt-Sigmaringen
Gartenstr. 15
72458 Albstadt

Derk Rembold

Technische Informatik

Hochschule
Albstadt-Sigmaringen
Gartenstr. 15
72458 Albstadt
E-Mail: rembold@hs-albsig.de

Bernd Stauß

Wirtschaftsinformatik

Hochschule
Albstadt-Sigmaringen
Johannesstr. 6
72458 Albstadt
E-Mail: staussb@hs-albsig.de

Stefan Schwarzkopf

Technische Informatik

42aaS GmbH
Vor dem Kreuzweg 17
7070 Tübingen
E-Mail: stefan.schwarzkopf@42aas.de

Mario Jacobi

Technische Informatik

Hochschule
Albstadt-Sigmaringen
Gartenstr. 15
72458 Albstadt

ABSTRACT

Im Rahmen des Forschungsprojekts KI-VISOPRO (Entwicklung KI-basierter virtueller Sensoren zur Prozesssteuerung) wurde ein Sensorsystem entwickelt, welches komplexe Prozessgrößen bestimmen kann, die mit herkömmlichen Sensoren nicht trivial messbar sind. Hierzu wurde ein Ansatz der künstlichen Intelligenz gewählt, der dem supervised Deep Learning zuzuordnen ist. Ein essenzieller Bestandteil des Projektes ist der Aufbau und die Inbetriebnahme eines Demonstrators zum strukturierten Sammeln der Prozess- und Sensordaten mittels eines Enterprise Resource Planning (ERP) Systems. Die ermittelten Prozessgrößen können dabei auch wieder in den Prozess und das ERP-System zurückgespielt und dort zur Prozesssteuerung oder der Optimierung von Prozessschritten und -parametern verwendet werden.

Motivation

Sollen physikalische Werte als Ziel- oder Leitgrößen in Prozessen verwendet werden, so müssen diese zuverlässig gemessen werden können. Dies erfordert oftmals teure und im Betrieb, Einbau und Wartung aufwendige Sensoren oder Messprozesse. Ein Beispiel hierfür ist der Messprozess zur Bestimmung der absoluten Feuchtigkeit in Baustoffproben oder Lebensmitteln wie Kaffeebohnen.

In beiden Fällen wird oftmals die Darr-Methode (gravimetrische Messung) eingesetzt. Dabei wird eine Probe gewogen, für eine bestimmte Zeit in eine Trockenkammer gestellt und letztendlich nochmals gewogen. Aus dem gemessenen Gewichtsverlust kann auf die anfangs enthaltene Feuchtigkeit geschlossen werden. Dieser aufwendige Messprozess benötigt einen manuellen Eingriff, welcher den Gesamt-Produktionsprozess verzögern kann.

Ein weiteres Beispiel findet man in der Bestimmung des Zuckergehalts während enzymatischer Umsetzungs-

bzw. Fermentationsprozesse, wie sie z.B. bei der Bierherstellung auftreten. Bei der Bestimmung des Zuckergehalts in Flüssigkeiten wird üblicherweise das Refraktometer oder vorzugsweise die Spindel (Saccharometer) verwendet. Die Spindel selbst zeigt dabei nicht direkt den Zuckergehalt an, sondern die Dichte der Flüssigkeit. Man weiß, dass die Dichte des reinen Wassers bei ca. 4°C 1000 kg/m³ beträgt. Der Zucker in der Flüssigkeit beeinflusst die Dichte der Lösung, woraus sich von der Dichtemessung auf den Zuckergehalt schließen lässt. Auch hier ergibt sich ein manueller Eingriff in den Prozessablauf, da Proben entnommen werden müssen. Die Spindel-Messung erscheint auf den ersten Blick verhältnismäßig unproblematisch, allerdings muss die Probe vor der Messung auf eine bestimmte Temperatur (meist 20°C, selten 70°C) gekühlt bzw. geheizt werden, da Saccharometer auf diese Temperaturen kalibriert sind. Hierdurch entsteht eine weitere zeitliche Verschiebung, bevor der Messwert als Prozessparameter verwendet werden kann.

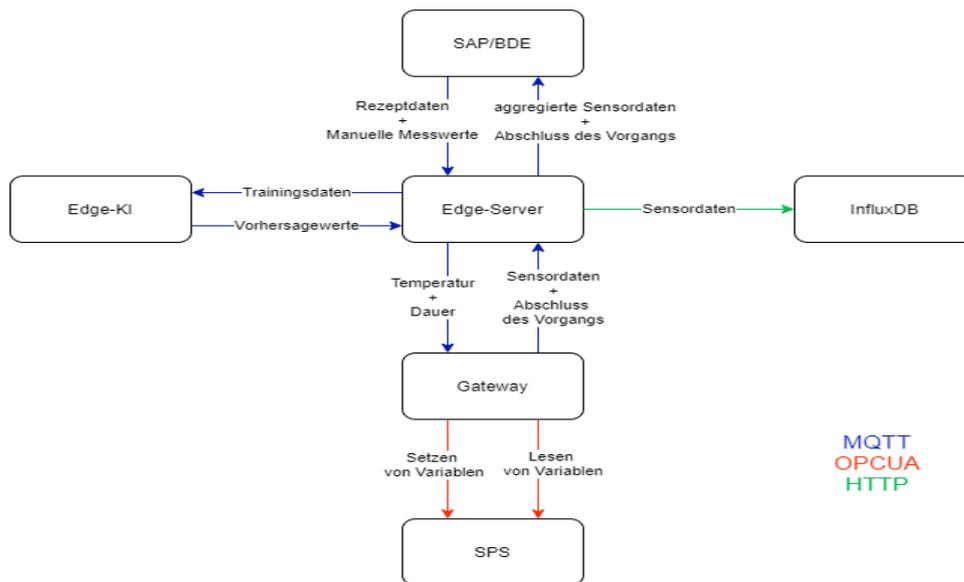


Abbildung 1: KI-VISOPRO vereinfachte Infrastruktur

Bei kleineren Temperaturunterschieden können Umrechnungstabellen helfen, jedoch erfordern diese ein sehr genaues Ablesen. Zudem können Störfaktoren wie Schwebstoffe oder Auftrieb durch Kohlensäure die Messung stark beeinflussen.

Für die Bestimmung des Zuckergehaltes gibt es bereits Geräte, die „inline“ Prüfungen, d.h. ohne explizite Probenahme, ermöglichen. Dabei wird unter anderem Ultraschall eingesetzt, um anhand von Laufzeitunterschieden und Schallabsorption auf die Dichte der Flüssigkeit zu schließen. Auch hier wird ähnlich wie bei der Spindel nicht direkt der Zuckergehalt, sondern die Dichte der Flüssigkeit gemessen. Diese Geräte sind jedoch sehr teuer und müssen regelmäßig von Fachpersonal gepflegt werden.

Im Rahmen des Projektes KI-VISOPRO wurden insbesondere die Messung des Zuckergehalts von Flüssigkeiten, am Beispiel Bierwürze, in den Mittelpunkt der Untersuchungen gerückt.

Beschreibung des Gesamtsystems

Abbildung 1 zeigt auf unterster Ebene den technischen Prozess, welcher im konkreten Fall auf einer Bierbrauanlage der Firma Speidel durchgeführt wird. Der technische Prozess wird über eine speicherprogrammierbare Steuerung (SPS) gesteuert. Über die SPS können einige Sensorwerte, z.B. Druckwerte, Stromverbrauchswerte der Aktoren und Temperaturwerte, ausgelesen werden. Zudem wird ein Sensormodul der Firma ifm verwendet, um das Abnehmen des Deckels des Bierbraukessels zu erkennen. Letztendlich gibt es noch eine Waage der Firma Kern, um das enthaltene Gewicht des Bierbraukessels zu messen. Die verschiedenen Komponenten werden in einem Gateway über unterschiedliche Protokolle und Technologien, bspw. OPC-UA, RS232 oder REST-Interface, gesammelt und zentral versendet. Durch dieses

Gateway wird zusätzlich die SPS von der Außenwelt getrennt, so dass von außen über das Internet nicht direkt auf die SPS zugegriffen werden kann. Zur Kommunikation mit der Außenwelt wird ein MQTT-Broker verwendet, der die Nachrichten an seine Subscriber verteilt. Zu diesen Subscribern gehört der Edge-Server, welcher die Nachrichten verarbeitet, die Kommunikation der Komponenten überwacht und die Sensorwerte in die InfluxDB Datenbank abspeichert. Der Edge-Server enthält eine State-Machine, welche den aktuellen Zustand eines aktiven Brauprozesses abbildet. Über die Nachrichten und die InfluxDB hat der Edge-KI Rechner Zugriff auf die Sensorwerte und Zustandsgrößen des technischen Prozesses und kann damit den Zuckergehalt über eine künstliche Intelligenz (KI) bestimmen. Auf oberster Ebene wird den Mitarbeitern über ein System zur Betriebsdatenerfassung (BDE) eine grafische Oberfläche (GUI) angeboten, über die eine Steuerung des Prozesses und eine Interaktion mit den übrigen Komponenten möglich wird.

Detaillierte Beschreibung des Gesamtsystems

Virtuelle Sensoren

Virtuelle Sensoren (VISO) bestehen aus unterschiedlichen physikalischen Sensoren und berechnen oftmals eine ansonsten schwer zu messende Größe. Dabei ist das Messen der notwendigen Sensordaten und die Wahl der passenden Messverfahren eine der wichtigsten und aufwendigsten Aufgaben. Im Abschnitt Motivation wurde bereits ein Beispiel erwähnt, in dem Ultraschall eingesetzt wird, um die Dichte einer Flüssigkeit zu messen um dann erst auf den Zuckergehalt zu schließen. Somit handelt es sich auch hier bereits um einen VISO, da die Zielgröße nicht direkt gemessen, sondern über verschiedene Sensoren und Algorithmen indirekt ermittelt wird. Zur Berechnung wird entweder ein physikalisches Modell,

ein analytisches Modell oder eine Kombination aus beiden eingesetzt. Diese Modelle sollen die physikalischen Zusammenhänge abbilden und anhand der Messwerte zuverlässige Ergebnisse liefern.

Physikalische Modelle werden hauptsächlich eingesetzt, wenn die Zusammenhänge der verschiedenen Zustandsgrößen bekannt sind und alle notwendigen Größen einfach gemessen werden können.

Analytische Modelle kommen oft zum Einsatz, wenn der genaue physikalische Zusammenhang noch nicht komplett bekannt ist oder bestimmte Größen nicht mit gewöhnlichen Sensoren gemessen werden können. In diesen Fällen kann dann eine KI eingesetzt werden, um aus den vorhandenen Daten eine Beziehung zur Zielgröße zu bestimmen und diese Vorhersage durch neue Daten stetig zu verbessern.

In der Regelungstechnik wird für diesen Zweck ein sogenannter Beobachter (als Beispiel für ein physikalisches Modell) verwendet [1]. Für eine genaue und passende Regelung werden verschiedene Zustandsgrößen benötigt, die aber bei vielen Systemen nur unter großem Aufwand messbar sind. Der Beobachter schafft es hier mit Hilfe eines mathematischen Modells (Differenzialgleichungen) die Zustandsgrößen rechnerisch zu bestimmen. Damit das Modell eine ausreichende Genauigkeit erreicht, werden Sensorwerte aus dem realen Prozess eingespielt. David Luenberger entwickelte dafür den Luenberger-Beobachter [2]. Ähnlich zum Luenberger-Beobachter beschreibt Gazaille in [3] einen Soft-Sensor um Größen aus einem Bioprozess zu ermitteln. Dieser besteht sowohl aus einem realen Sensor, als auch aus einem Schätzer, um die gewünschte Größe zu ermitteln.

Wolfgang Samenfin [4] patentierte ein Verfahren, um die Qualität, insbesondere die Viskosität, von Öl bei Verbrennungsmotoren zu bestimmen. Er zeigte über Schaubilder einen Zusammenhang zwischen der Viskosität und der Verunreinigung. Dabei werden bei dem vorgeschlagenen Verfahren lediglich die Temperatur und der Öldruck gemessen. Beides sind Größen, die sich mit herkömmlichen und einfachen Sensoren bestimmen lassen. So lässt sich über eine mathematische Funktion mit Temperatur und Öldruck als Parameter, die Viskosität bestimmen. Ein weiterer interessanter Aspekt ist der Einsatz von Kontextdaten, bspw. die Zeitpunkte der Serviceintervalle und des Ölwechsels. Diese werden berücksichtigt und beeinflussen das Datenmodell. Dieser Ansatz fließt mit in das Projekt ein und wird weiterverfolgt, um bspw. Unterschiede in den verwendeten Rohstoffen festzustellen und diese Erkenntnisse an das Modell weiterzugeben.

Rolle des Edge-KI Rechners für den virtuellen Sensor

Viele physikalische Größen können mit herkömmlichen und einfachen Sensoren nur bestimmt werden, wenn die physikalischen Zusammenhänge des technischen Prozesses bekannt sind. Ein bereits erwähntes Beispiel ist die Bestimmung des Zuckergehalts in Bierwürze. Ein einfaches Messprinzip dabei ist die Bestimmung der Dichte mit zwei Drucksensoren, die mit einem bekannten Höhenunterschied verbaut sind (siehe Abbildung 2). Drucksensoren sind auf dem Markt gut erhältlich und können

ohne allzu großen Aufwand in den Braubehälter installiert werden. Die Dichte kann jedoch nicht direkt über den hydrostatischen Druck bestimmt werden, da weitere Parameter in die Dichteberechnung miteinfließen. Ein Beispiel ist die geometrische Anordnung der Drucksensoren, insbesondere die Entfernung zur Heizspirale im Bierbraukessel. Auch kann die Temperaturverteilung innerhalb der Flüssigkeit variieren, wodurch sich deutliche Schwankungen in der Dichtemessung ergeben können. Dies wurde in den ersten Vorversuchen des Projektes als Problem identifiziert und durch eine geschickte Anordnung der Sensoren weitestgehend gelöst.



Abbildung 2: Zwei Ein-Zoll-Einschraubstutzen für Drucksensoren

Trotzdem ist es wichtig nicht nur die Sensorwerte (Druck, Temperatur), sondern auch weitere Zustandsgrößen (Heizspirale an/aus) zu erfassen. Dadurch wird die Modellbildung mit einem physikalischen Modell jedoch schwierig, da die Gesamtheit aller Zustandsgrößen das physikalische Modell sehr komplex gestalten würde. Aus diesem Grund wird vorgeschlagen, das physikalische Modell durch eine KI zu ersetzen.

Oftmals sind herkömmliche Rechner für diese Aufgaben nur bedingt geeignet, da das Trainieren der KI sehr rechenintensiv sein kann. Dabei sind vor allem eine leistungsfähige CPU oder eine für KI geeignete GPU sehr wichtig. Das Trainieren der KI wird daher auf einen Rechner, im Weiteren Edge-KI genannt, mit dedizierter GPU ausgelagert.

Sensorwerte aus einem Wassertest-Experiment zur Validierung der Dichtemessung

Um das Messverfahren zu validieren, wurde ein Versuch ohne Malz durchgeführt. Abbildung 3 zeigt den Temperaturverlauf des Wassers im Braukessel. Zunächst steigt die Temperatur auf 63°C und wird dort für ca. 70 Minuten gehalten. Diese Phase nennt man in der Brautechnik die erste Rast (rot). Das Halten der Temperatur wird über einen Zweipunktregler der SPS gesteuert. Danach erfolgt

ein weiterer Temperaturanstieg auf 73°C und das Halten der Temperatur für ca. 5 Minuten. Dieser Abschnitt wird als zweite Rast (blau) bezeichnet. Danach erfolgt mit dem Abmaischen (grün), im Diagramm und im Weiteren als Rast 3 bezeichnet, der letzte Temperaturanstieg auf 78°C mit einer Haltedauer von 5 Minuten.



Abbildung 3: Temperaturverlauf von Wasser

Abbildung 4 veranschaulicht den Verlauf der Dichte, welche durch die zwei Drucksensoren berechnet wurde. Aufgrund des Temperaturanstiegs, sinkt auch die Dichte des Wassers (orange). Sobald die Temperatur von der SPS gehalten wird, bleibt die Dichte des Wassers im Mittel konstant (rot). Beim Aufheizen zur zweiten Rast nimmt die Dichte des Wassers weiter ab (gelb), bis die zweite Rast erreicht wird (blau). Beim Aufheizen zur dritten Rast nimmt die Dichte des Wassers nochmals ab (lila) und wird wieder konstant, solange die Rast anhält (grün). Der dazugehörige Temperaturverlauf ist in Abbildung 3 abgebildet.



Abbildung 4: Verlauf der Dichte

Sensorwerte aus einem Maische-Prozess zur Validierung der Dichtemessung

Im Folgenden wird auf die Sensorwerte während des Brauprozesses eingegangen, insbesondere auf die Werte der zwei Drucksensoren. Hierbei wird der Bierbraukessel mit Wasser gefüllt und auf 40°C aufgeheizt. Dann wird Malz zugegeben und der Inhalt auf 63°C erhitzt und diese Temperatur für ca. 70 Minuten gehalten. Abbildung 5 zeigt diesen Temperaturverlauf. Ziel dieser ersten Rast ist es, einen biochemischen Prozess in Gang zu setzen. Dabei wandeln im Malz enthaltene Enzyme (β -Amylase) die Stärkemoleküle des Malzes hauptsächlich in kurzketige und vergärbare Zucker (Maltose) um.



Abbildung 5: Temperaturverlauf von Maische

Abbildung 6 zeigt den geglätteten Dichteverlauf (grün), berechnet aus der Differenz der beiden Drucksensoren. Die Druckdifferenz nimmt während der ersten Rast stets zu. Weitere Verläufe in Abbildung 6 sind die manuellen Zuckermessungen mit der Bierspindel. Hier werden Proben aus dem Bierkessel entnommen und die Bierspindelwerte direkt nach Entnahme (unkalibriert, hellblau) und nach dem Aufheizen (kalibriert, orange) aufgezeichnet. Hier ist zu sehen, dass die Werte der Dichte und der manuellen Messungen ähnliche Verläufe haben. Nach 70 Minuten geht der Prozess in seine nächste Phase, die zweite Rast. Es gibt hier ein Aufheizen auf 73°C und ein Halten für 5 Minuten. Dabei werden verbleibende Stärkemoleküle durch die Enzyme (α -Amylase) in größtenteils langkettige, nicht vergärbare Zucker (Dextrin) umgewandelt. Durch das Aufheizen der Flüssigkeit ist in Abbildung 6 auch eine Abnahme der Dichte ersichtlich. Nach Ende der zweiten Rast gibt es ein letztes Aufheizen auf 78°C. Bei dieser Rast werden die Enzyme und Proteine durch die Hitze zerstört und somit der biochemische Prozess zunächst beendet. Die letzte Phase des Prozesses wird Läutern genannt. Hier wird das Malz aus der Bierwürze gehoben und anschließend Wasser über den Trester (Überbleibsel vom Malz) gegossen. Das Wasser löst den noch im Trester enthaltenen Zucker. Bei den manuellen Messungen ist hierbei eine Abnahme der Bierspindel-Werte zu beobachten.



Abbildung 6: Verlauf der Dichte und Zuckergehalts

Datenmanagement

Für das Trainieren des virtuellen Sensors wird eine große Datenmenge benötigt. Dabei ist zunächst zu bestimmen, welche Daten im Prozess eine besondere Bedeutung haben. Dies erfolgt i.d.R. durch eine Kombination aus Expertenwissen und experimentellen Analysen. Letzteres kann zum einen durch das automatisierte Aufzeichnen von Daten bei der Prozessausführung geschehen. Zum anderen können aktiv Experimente für spezifische Situationen geplant werden. Der zweite Ansatz ist insbesondere dann sinnvoll, wenn lediglich Teile des Gesamtprozesses für das Training der Modelle relevant sind. Der Gesamtprozess zur Herstellung von Bier beinhaltet nach der Herstellung der Bierwürze noch weitere Schritte, z.B. das Hopfenkochen oder Reifen. Diese Schritte sind aber für die Datengewinnung völlig unerheblich, weshalb auf eine Weiterverarbeitung der Würze verzichtet wurde. Neben der erwarteten Zeitersparnis war bei der Entscheidung über die Datengewinnung die größere Flexibilität bei der Variation der Trainingsbedingungen zentral. Für die Generierung valider Trainingsdaten wurden drei wesentliche Anforderungen definiert:

1. Größtmögliche Transparenz über den Prozess und das Zustandekommen der Daten
2. Vermeidung von Fehlern bei der Versuchsdurchführung
3. Berücksichtigung von Kontextdaten, bspw. Prozessinformationen, Beschaffenheit der Rohstoffe etc.

Aus den Überlegungen zur letzteren Anforderung entstand die Annahme, dass Umweltbedingungen (bspw. Umgebungstemperatur und Luftfeuchtigkeit) oder die Beschaffenheit der Ausgangsstoffe Einfluss auf die Versuche haben können. Dies ist auch unmittelbar einleuchtend,

Vorgangsnummer	Arbeitsplatz	Werk	Steuerschlüssel	Beschreibung	Rüstzeit	Maschinenzeit	Personalzeit	Rüsten variable
0005	P01-001	DF01	Z001	PH-Messer kalibrieren (manuell)	5	0	0	0
0010	P01-001	DF01	Z002	Wasser einfüllen (manuell)	5	0	0	0
0015	P01-001	DF01	Z004	Temperieren	0	0	0	5
0020	P01-001	DF01	Z002	PH-Wert messen und anpassen (manuell)	3	0	0	0
0025	P01-001	DF01	Z002	Malz zugeben (manuell)	2	0	0	0
0030	P01-001	DF01	Z004	Heizen auf Zieltemperatur	0	0	0	5
0035	P01-001	DF01	Z004	Rast 1	0	70	0	0
0045	P01-001	DF01	Z004	Heizen auf Zieltemperatur	0	0	0	5
0050	P01-001	DF01	Z004	Rast 2	0	5	0	0
0065	P01-001	DF01	Z004	Heizen auf Zieltemperatur	0	0	0	5
0075	P01-001	DF01	Z004	Rast 3	0	5	0	0
0085	P01-001	DF01	Z001	Treber heben (manuell)	5	0	0	0
0090	P01-001	DF01	Z002	Zuckergehalt messen (manuell)	3	0	0	0
0095	P01-001	DF01	Z002	Läutern (manuell)	10	0	0	0
0100	P01-001	DF01	Z005	Mengenfeststellung / Auftragsabschluss	1	0	0	0

Tabelle 1: Arbeitsplan für einen Brauversuch

tend, da erstens bei den Rohstoffen unterschiedliche Sorten verwendet werden und zweitens selbst bei der gleichen Sorte nur bedingt kontrollierbare Unterschiede in deren Beschaffenheit wie z.B. Mehlgehalt, Wassergehalt etc. zwischen zwei Losen vorkommen können. Die chemisch-physikalischen Eigenschaften des Malzes liegen der Lieferung jeweils in Form eines Laborberichts bei und können so für das Training der KI verwendet werden. Hierfür sind eine entsprechende Erfassung, Verwaltung und Bereitstellung der Daten notwendig.

Beim Sammeln von Trainingsdaten gibt es verschiedene Fehlerquellen. Eine Quelle von Fehlern besteht in der nicht sachgemäßen Ausführung der Versuche, also z.B. dem Überspringen oder Auslassen von Prozessschritten, der Missachtung von Prozessvorgaben etc. Somit muss der Versuchsaufbau diese Einflüsse so gut es geht auf ein Mindestmaß reduzieren und gegebenenfalls bei nicht konformer Prozessausführung dies entsprechend vermerken. Eine weitere Fehlerquelle ist die manuelle Erfassung von Messergebnissen. Diese falschen Messergebnisse können zum einen durch falsches Ablesen der Messinstrumente oder zum anderen durch Eingabefehler bei der Messwertaufzeichnung entstehen. Auch ein einfaches Auslassen von einzelnen Messintervallen ist möglich, wenn der Labormitarbeiter die Durchführung der Messung vergisst.

Abschließend ist die Nachvollziehbarkeit der erfassten Daten und die Möglichkeit der Prozessüberwachung gerade bei unterschiedlichen Rezepten ohne eine systematische Verwaltung dieser Daten über ein betriebliches Informationssystem (bIS) nur schwer umsetzbar.

Im Bereich bIS besteht bei den Projektteilnehmern eine langjährige Erfahrung mit ERP-Systemen, speziell mit

dem Produkt S/4HANA von SAP. Dieses steht der Hochschule Albstadt-Sigmaringen als eigenes System zu Ausbildungs- und Forschungszwecken zur Verfügung. Teile der Anforderungen können bereits mit der Standardauslieferung der Systems abgedeckt werden, weitere Funktionen können aufgrund der Flexibilität des Systems mit vertretbarem Aufwand durch Systemkonfigurationen und Code Anpassungen hinzugefügt werden.

Durch das große Ökosystem an verfügbaren Erweiterungen und Add-Ons, rund um S/4HANA, kann die Standardsoftware an spezifische Rahmenbedingungen angepasst werden.

Die Führung der Labormitarbeiter durch den Versuch ist ein zentraler Baustein des bIS. Zunächst müssen hierfür

sämtliche relevanten Daten identifiziert und eine passende SAP Darstellung gewählt werden. Im Wesentlichen werden die Produktstruktur, also die Ausgangsstoffe mit entsprechenden Mengenangaben, sowie der konkrete Ablauf des Versuchs mit diesen Daten beschrieben. Jedem Vorgang können Attribute beziehungsweise Merkmale zugeordnet werden, die entweder Vorgaben für die Prozessautomatisierung darstellen oder für die Messdatenerfassung verwendet werden.

Die Realisierung in SAP erfolgt mittels Arbeitsplänen, siehe Tabelle 1, und den einzelnen Vorgängen zugeordneten Prüfmerkmalen. Die Prüfmerkmale können sowohl Prozessvorgaben als auch zu erfassende Messgrößen darstellen. Die Unterscheidung erfolgt über spezifische Parameter in den jeweiligen Stammprüfmerkmalen. Die Erfassung der Prüfergebnisse wird in SAP mittels Prüfpunkten umgesetzt, so dass seine den Versuch begleitende Prüfung ermöglicht wird.

Wie bereits erwähnt, gibt es natürliche Schwankungen in der Zusammensetzung des Hauptrohstoffs Malz, so dass diese Informationen zu jedem Sack des Malzes gespeichert werden müssen. Dies erfordert eine konsequente Identifikation der Malzsäcke, die als Charge im bIS verwaltet werden, sowie eine konsistente Erfassung der in den Versuchen eingesetzten Chargen. SAP bietet für dieses Szenario mit der Chargenklassifikation ein sehr einfach zu verwendendes aber auch mächtiges Instrument zur Beschreibung von Chargeneigenschaften. Relevante Daten, wie Wasser-, Mehlgehalt etc. werden aus dem Begleitpapier (Laborbericht) als Chargenmerkmal gespeichert, siehe Tabelle 2. Die manuelle Eingabe der Daten

kann durch den Einsatz einer SAP-Erweiterung, die in einem parallel durchgeführten Studierendenprojekt entstand, und auf Optical Character Recognition (OCR) aufbaut, einfach automatisiert werden. Hierdurch kann eine hohe Datenqualität erreicht werden und jede Anlieferung führt direkt zu einem Wareneingang mit einer neuen, automatisch klassifizierten Charge.

Parameter	Einheit	Messwert
Wassergehalt	%	3,5
Extrakt Mehl lufttrocken	%	79,4
Extrakt Mehl wasserfrei	%	82,2
Sortierung 2,8mm	%	81,9
Etc.	Etc.	Etc.

Tabelle 2: Ausschnitt aus Begleitpapier (Laborbericht)

Jeder Versuch wird über einen eigenen Auftrag im bis abgebildet. Zum Auftrag werden zunächst die eingesetzten Rohstoffe, deren Menge und die ausgewählte Charge

Sensordaten von SPS empfangen, alle 10 Sekunden						
Topic	QoS	Mqtt Payload	Sensorname	Datentyp	Einheit	Anmerkung
sps/edge/values	2	{"timestamp":<Long>,"operation":<String>,"batchid":<String>,"emergencyshutdown":<bool>,"hatch":<bool>,"sensorValues":{"pressure_high":20.7,"pressure_low":31.65,"temperature":68.3,"weight":34.8,"<String>":<Float>...}}	temperature	Float	°C	
			weight	Float	kg	-100 wenn Waage aus ist
			pressure_high	Float	mbar	
			pressure_low	Float	mbar	
			power_str	Float	mA	
			voltage_str	Float	V	
			power_heat	Float	mA	
			voltage_heat	Float	V	
			emergency_shutdown	Boolean		True ist der kritische Fall, False ist der normale Fall
			hatch	Boolean		False nicht drauf und True drauf
			timestamp	EpochTime(Long)		

Tabelle 3: Beispiel für eine MQTT-Nachricht (Toggle-Nachricht)

erfasst und über eine Bestandsbuchung, also einer Warenentnahme, dem Auftrag zugeordnet.

Der Auftrag im bis stellt somit das zentrale Element bei der Bereitstellung der Prozessvorgaben, Kontextdaten sowie für die Durchführung der Messungen dar. Diese Daten müssen auch für die anderen Komponenten des Gesamtsystems verfügbar sein. Grundsätzlich stehen für den systemübergreifenden Austausch von Daten ins S/HANA verschiedene Ansätze zur Wahl, bspw. IDoc, RFC, WebServices, OData-Service etc.

Bereits in zahlreichen Projekten hat sich die Kommunikation über das im Internet of Things (IoT) Bereich weitverbreitete MQTT-Protokoll bewährt.¹

Waren in der Vergangenheit für diese Kommunikation noch separate Komponenten wie z.B. SAP JCo, SAP PCo oder andere Middleware, zwischen SAP und der übrigen Infrastruktur notwendig, so kann diese nun direkt integriert über eine SAP Klasse erfolgen. Ein permanentes Überwachen der abonnierten Topics wird mit einem SAP-Daemon realisiert.

Gemeinsam mit dem Projektpartner top flow wurde eine in das ERP-System vollständig integrierte MES-Lösung implementiert. In großen Teilen konnte auf Standardkomponenten der Firma top flow zurückgegriffen werden und durch das flexible Framework auf die Projektspezifika angepasst werden. Das Kernstück der Erweiterung besteht aus einer BDE-Anwendung, über das Labormit-

arbeiter alle relevanten Daten angezeigt und manuelle Messdaten eingetragen werden können.

Systemarchitektur und -betrieb

Der schematische Aufbau des Gesamtsystems wurde bereits vorgestellt. Im Folgenden soll nun näher auf die Orchestrierung der einzelnen Komponenten eingegangen werden.

Grundlage für die Kommunikation sämtlicher Teile des Gesamtsystems ist das MQTT-Protokoll. Über einen Broker können die Teilsysteme auf verschiedenen Topics miteinander kommunizieren. Ein Beispiel für eine der Nachrichten ist in Tabelle 3 zu sehen. Zur Benennung der Topics wurde die Konvention "Sender/Empfänger/ Bedeutung" verwendet. An dem Topic edge/sps/toggle erkennt man bspw., dass es sich um eine Toggle-Nachricht

handelt, die vom Edge-Server verschickt und von der SPS, bzw. vom Gateway, verarbeitet wird.

Zur Beschreibung der Datenflüsse im Rahmen der Prozessausführung wurde ein vereinfachtes Sequenzdiagramm verwendet, das in Abbildung 7 dargestellt ist.

Aus der Darstellung erkennt man, dass die Informationen über den Edge-Server in zwei Richtungen weitergeleitet werden. Zum einen werden die Daten in einer InfluxDB-Datenbank gespeichert und dort für das Training der neuronalen Netze bereitgestellt. Zum anderen werden bestimmte Daten an die SPS durchgeschleift. Zusätzlich wird auf dem Edge-Server der aktuelle Prozesszustand gespeichert und verwaltet.

Während der Design- und Testphase erwiesen sich im Hinblick auf die Prozessausführung immer wieder Ausfälle einzelner Komponenten im Gesamtsystem als problematisch. So war bspw. zunächst beim Ausfall des MQTT-Brokers nicht gewährleistet, dass keine Daten verloren gehen oder bei einem Ausfall des Edge-Servers die erfassten Daten dennoch in die Datenbank geschrieben werden. Eine Risikoanalyse ergab, dass insbesondere die folgenden Fehlerfälle ein hohes Risiko für Daten- bzw. Konsistenzverlust innerhalb der Prozessausführung darstellen und aus diesen Gründen eine Risikovermeidung notwendig ist. Eine Mindmap über verschiedene Risikoszenarien ist in Abbildung 8 dargestellt.

ERP-System aus. Konkret erfolgt unter Verwendung des Java-Connectors der Aufruf bestimmter RFC-fähiger Bausteine im S/4HANA System.

¹ Als Middleware fungiert eine JAVA-Application. Diese subscribt sich als MQTT-Client auf die zur Anlagensteuerung relevanten Topics, verarbeitet und interpretiert die Daten und löst schließlich Aktionen im

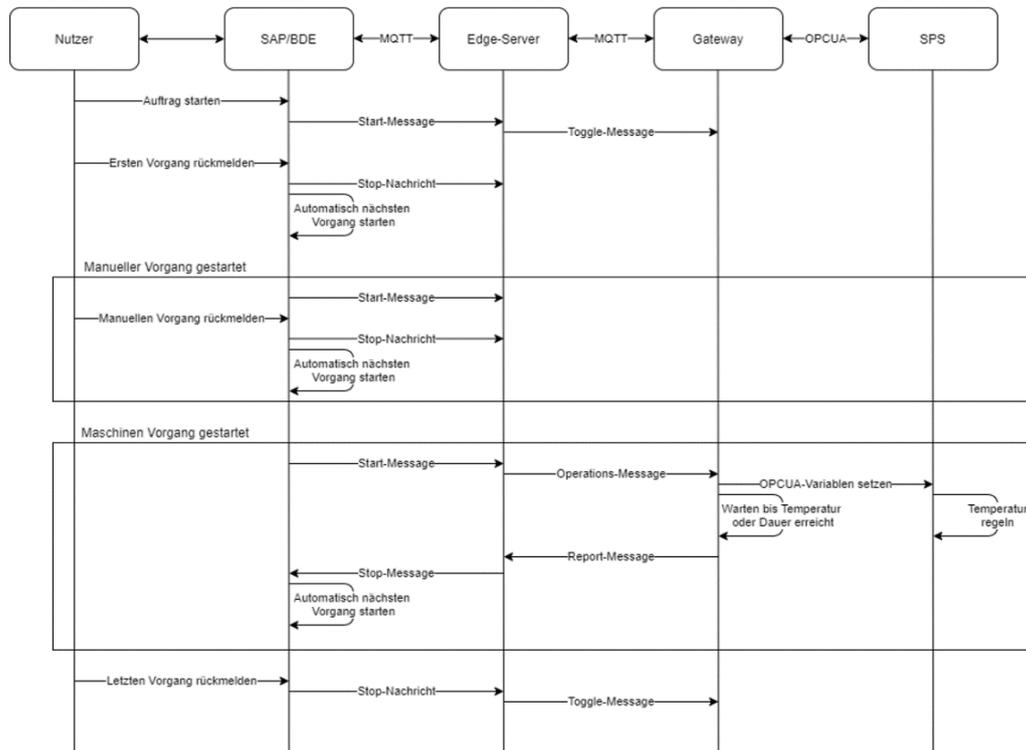


Abbildung 7: Sequenzdiagramm

Folgende Hauptfehlerquellen wurden identifiziert:

- Absturz Gateway
- Absturz State-Machine
- Absturz MQTT-Broker
- Absturz SAP MQTT-Daemon

Der erneute Start nach dem Absturz einer der obigen Komponenten erfordert komponentenspezifische Lösungen, welche im Folgenden näher erläutert werden.

Um den grundsätzlichen Ausfall von Komponenten feststellen zu können, verschicken SAP und Edge-Server in regelmäßigen Abständen eine Alive-Nachricht. Für das Gateway übernimmt die Nachricht zur Übermittlung der Sensorwerte diese Funktion, da diese bereits in regelmäßigen Abständen versendet wird.

Beim Gateway werden einige Zustandsinformationen (Zeitstempel, Dauer, Temperatur und Status) beim Empfangen auf die SPS geschrieben und dort gespeichert. Beim Neustart werden diese Informationen eingelesen und von der SPS beim Restart verwendet.

Nach dem Absturz wird der Edge-Server automatisch neu gestartet. Der aktuelle Zustand kann anhand von SAP Alive-Nachrichten ermittelt und die State-Machine wieder in den passenden Zustand gebracht werden.

Damit bei einem Ausfall des Internets oder des zentralen MQTT-Brokers trotzdem Daten von den Sensoren nicht verloren gehen, wurde ein lokaler MQTT-Broker installiert, der die Nachrichten im Fehlerfall zwischenspeichert. Sobald die Verbindung zum zentralen MQTT-Broker wieder hergestellt ist, werden die Daten an diesen weiter-

geleitet und in die zentrale Datenbank geschrieben. Zusätzlich werden die Daten lokal auf dem Gateway in eine separate InfluxDB geschrieben, wodurch selbst bei einem Internetausfall die Daten weiterhin vorhanden sind oder bei Übertragungs- oder Nutzerfehlern ein Backup vorhanden ist. In der Versuchsdurchführung hat sich dieses Vorgehen als sinnvoll erwiesen, da sowohl ein Absturz der lokalen Datenbank als auch ein Internetausfall auftraten, dabei aber keine Prozessdaten verloren gingen. Auf Seiten von SAP wurde ein Service eingerichtet, der den MQTT-Daemon überwacht und diesen neu startet, falls die Verbindung zum Broker verloren geht.

Modellbildung und Training

Sämtliche Sensordaten aus dem Bierbrauprozess werden in einer InfluxDB-Datenbank gespeichert. InfluxDB ist eine Zeitreihen-basierte Datenbank, deren Fokus auf der Verarbeitung von Daten mit Zeitinformationen liegt. Aus diesem Grund ist sie sehr gut geeignet, um bspw. Sensordaten zu speichern und diese aufzubereiten. Es werden jeweils von sieben Sensoren Werte gesammelt: Temperatursensor, Drucksensor-oben, Drucksensor-unten, Deckel-Sensor, Leistung-Heizung, Leistung-Rührer und Gewicht. Aus diesen Sensorwerten können weitere Daten abgeleitet werden, bspw. enthält die Leistung der Heizung Informationen über deren aktuellen Zustand (Heizung an oder aus). Das Gateway sendet etwa alle zehn Sekunden eine Nachricht mit den aktuellen Sensordaten. Neben diesen beobachtenden Daten gibt es auch manu-

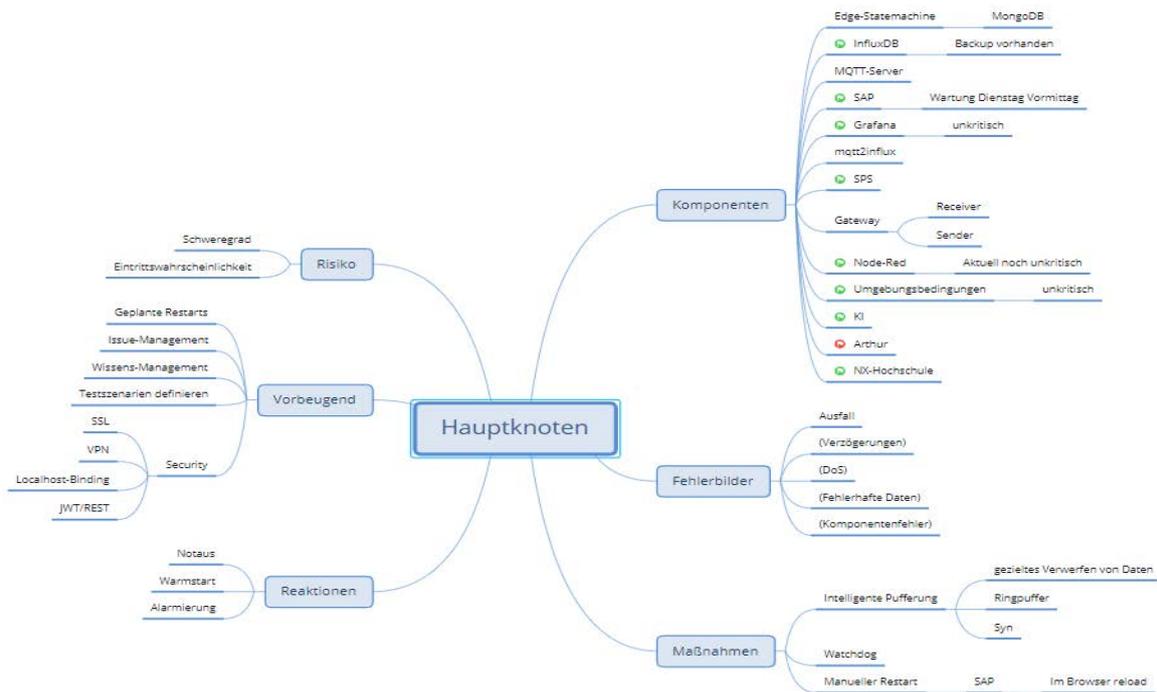


Abbildung 8: Mindmap zur Risikobewertung

elle Messungen. Dazu zählt unter anderem das Ziel-Feature (Target-Feature), der Bierspindelwert. Weitere manuelle Messungen sind bspw. der pH-Wert und die Messung des Bierspindelwerts bei einer auf 70°C aufgeheizten Probe. Diese Messungen benötigen einen manuellen Eingriff in den Prozess und werden etwa alle zehn Minuten getätigt, über das bIS eingetragen und in die Datenbank übernommen. Aufgrund der deutlich geringeren Frequenz der manuellen Messungen müssen diese Datensätze vor dem Training noch aufbereitet werden. Hierfür werden die Daten unter anderem interpoliert, um jedem Zeitpunkt einen Zielwert zuweisen zu können. Die Ergebnisse der Aufbereitung werden in eine CSV-Datei gespeichert, um die weitere Verarbeitung zu erleichtern und Zeit für zusätzliche InfluxDB-Queries zu sparen.

Während des Trainierens des Netzes hat es sich als vorteilhaft erwiesen, bestimmte Messwerte nicht als direktes Eingangs-Feature zu verwenden. Beispielsweise haben Stufenfunktionen einen sehr starken Einfluss auf die Vorhersagen des Netzes, da bei der zeitlichen Ableitung durch die Sprünge sehr große Werte entstehen, die dann einen ebenso großen Einfluss auf die Optimierung der Gewichte haben. Diskrete Zustandsgrößen werden jedoch als Hilfsgrößen verwendet, um bspw. das Öffnen und Schließen des Bierbraudeckels zu dokumentieren und Messungen, die zu diesen Zeitpunkten erfolgten herauszufiltern. Aus den gesammelten Sensorwerten werden im aktuellen Modell nur noch drei verwendet, die Temperatur, der obere Druck sowie die Druckdifferenz. Als Zielgröße (Target-Feature) dient der unkalibrierte Zuckermesswert.

Für zeitbasierte Daten werden oft LSTM-Modelle (Long-Short-Term-Memory) verwendet, um den zeitlichen Zusammenhang der Daten zu berücksichtigen. Diese zeichnen sich durch das Speichern bestimmter Zustandsgrößen (Hidden und Cell State) aus, wodurch zeitliche und dynamische Aspekte durch die Daten einen Einfluss auf die aktuelle Berechnung haben. Während des Trainings werden die Gewichte des Modells so angepasst, dass eine Fehlerfunktion minimiert und damit die Vorhersagen sich immer mehr den gemessenen Werten annähern. Das verwendete Modell besteht aus drei Schichten. Die erste Schicht besteht aus 32 LSTM Einheiten, darauf folgt eine Dense-Schicht mit 64 Einheiten und eine letzte Dense-Schicht mit nur einem Neuron, siehe Abbildung 9.

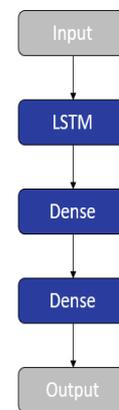


Abbildung 9: verwendetes LSTM

Die Verbindungen zwischen den einzelnen Schichten bzw. Einheiten werden Gewichte genannt und werden im Trainingsverlauf optimiert. Das gewählte Netz ist relativ

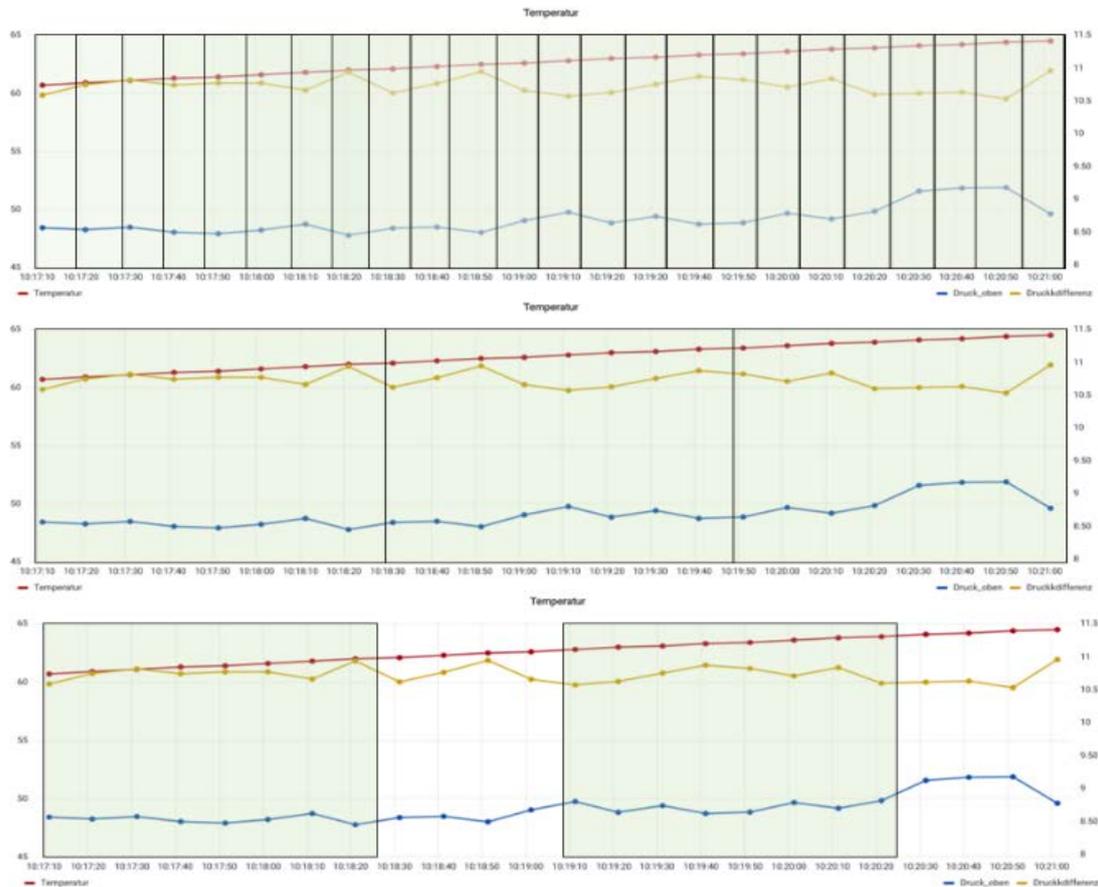


Abbildung 10: Optionen für die Festlegung von Zeitfenstern

klein, wobei beim Training auch größere Netze getestet wurden, durch die allerdings keine Verbesserung bzgl. der Vorhersagegüte erzielt werden konnte. Durch das kleine Netz können sehr flexibel und schnell verschiedene Trainingseinstellungen (Hyperparameter) verändert und getestet werden.

Bevor das Netz die Daten erhält, werden die Daten normalisiert, also in einen bestimmten Wertebereich transformiert, damit große und kleine Datenwerte den gleichen Einfluss haben. Zum Normalisieren wurde der Z-Score [5] verwendet.

Während des Trainings liest der Trainingsalgorithmus acht zeitlich aufeinanderfolgende Zeitpunkte mit den jeweiligen Attributen aus und fügt sie zu einer 8×3 Matrix zusammen. Die entstandene Matrix kann man sich als Zeitfenster im Diagramm vorstellen. Nun hat man mehrere Optionen, dieses Zeitfenster im nächsten Schritt zu verschieben (siehe Abbildung 10). Option eins (obere Graphik) ist das Zeitfenster um eine Zeiteinheit nach vorne zu schieben, sodass sich das neue und das alte Zeitfenster überschneiden. Eine zweite Option (mittlere Darstellung) ist, genau die Länge des Zeitfensters nach vorne zu gehen, so dass die Zeitfenster aneinander angrenzen, sich aber nicht überschneiden. Eine dritte Option (untere Graphik) ist es, einen größeren Schritt zu machen, so dass eine Lücke zwischen zwei Zeitfenstern entsteht, hier würden allerdings weniger Trainingsdaten entstehen.

Für das Netz wurden Optionen eins und zwei getestet. Im Training hat sich gezeigt, dass Option zwei, aneinander angrenzende, nicht überlappende Zeitfenster, bessere Ergebnisse liefern und die Modelle schneller konvergieren. Auf Option drei wurde aufgrund der sich verringernenden Datenmenge verzichtet.

Abbildungen 11 und 12 zeigen Versuche mit Testdaten an denen die prädiktive Validität des Modells überprüft

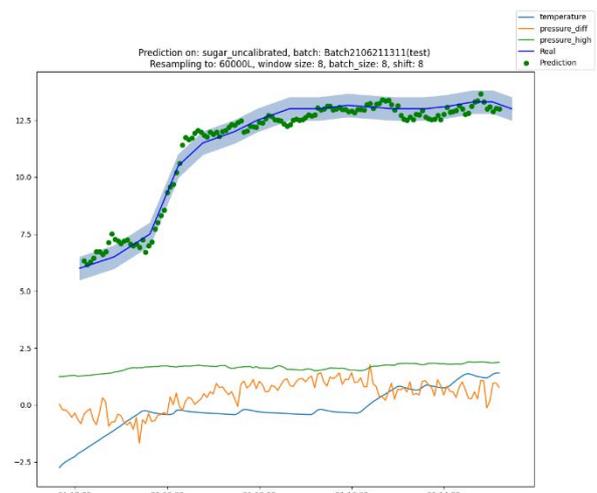


Abbildung 11: Testdaten, kleiner Fehler

wurde und bei denen das Netz sehr erfolgreich den Zuckerwert vorhersagt. Testdaten sind Daten auf die das Modell nicht trainiert wurde. Es ist klar zu erkennen, dass das Netz dem realen Verlauf folgt und die Vorhersagewerte größtenteils in dem blau eingezeichneten Toleranzbereich von 0,5 liegen.

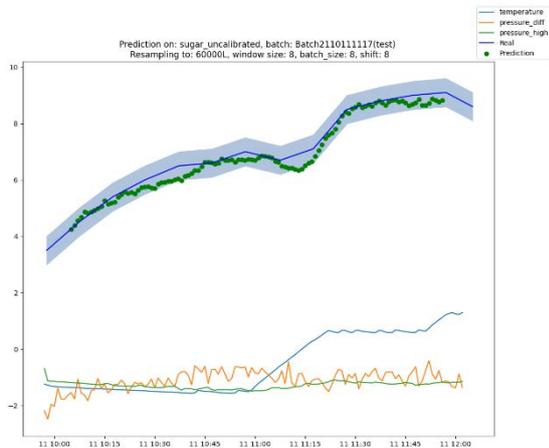


Abbildung 12: Testdaten kleiner Fehler

Bei einigen Versuchen liefert die KI jedoch auch schlechtere Ergebnisse. So ist in Abbildung 13 zu erkennen, dass die KI zunächst dem Zuckerverlauf gut folgt, dann jedoch deutlich abweicht und sich erst gegen Ende wieder den realen Werten annähert.

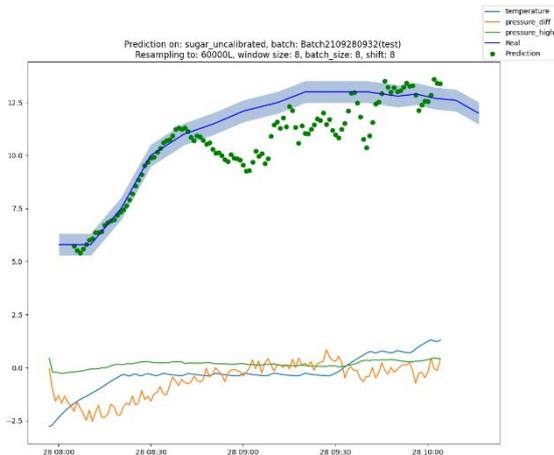


Abbildung 13: Testdaten, deutliche Abweichung

In Abbildung 14 ist ein weiterer Versuch abgebildet. Hier sind deutliche Ausreißer, die den Toleranzbereich verlassen, zu erkennen. Die KI kann grundsätzlich dem realen Verlauf trotzdem gut folgen und endet vor dem Läutern auf einem vergleichbaren Wert.

Während eines laufenden Versuchs kann die KI über Inferenz der Sensordaten live einen Vorhersagewert bilden. Dieser wird wiederum in InfluxDB geschrieben und kann anschließend live mit Hilfe von Grafana visualisiert und angezeigt werden, siehe Abbildung 15.

Das Netz braucht in diesem live Setup in fast allen Fällen weniger als eine Sekunde um das Ergebnis zu berechnen (Inferenzzeit), siehe Abbildung 16. Dadurch kann der

Vorhersagewert sehr schnell zurück in den Prozess zurückgespielt und dort weiterverarbeitet werden. Selbst die großen Ausreißer bleiben unter zwei Sekunden Inferenzzeit.

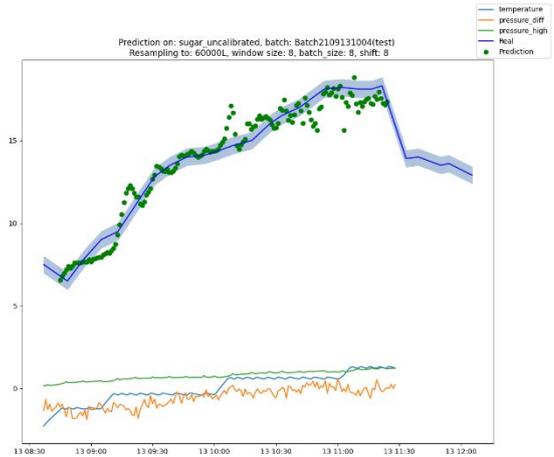


Abbildung 14: untrainierter Versuch mit Ausreißern



Abbildung 15: Vorhergesagter Zucker während eines Versuches

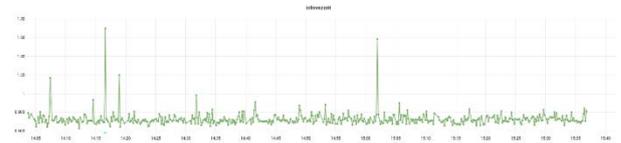


Abbildung 16: Inferenzzeit der Vorhersagewerte

Fazit

Durch die Prozessführung durch SAP/BDE und der Darstellung aller relevanten Informationen kann der Prozess und somit die Versuche deutlich einfacher und prozesssicherer durchgeführt werden. Ein zusätzlicher Effekt ist die automatische Dokumentation von möglichen Fehlern und allgemeinen Prozess- und Kontextinformationen, wodurch eine Prozessoptimierung ermöglicht wird.

Im Projekt waren verschiedene Herausforderungen zu lösen, so hatte bspw. die anfängliche Position der Drucksensoren zu einem erheblichen Temperatur- und damit auch Dichteunterschied geführt. Eine weitere Herausforderung stellte die Resilienz des Prozesses dar. In Rasten wird die Temperatur nur für eine gewisse Dauer gehalten, ein langer Ausfall der Komponenten kann durch Überschreiten dieser Prozesszeiten zum kompletten Versuchsabbruch führen. Kürzere Ausfälle, bis etwa 10 Minuten, kann das System verkraften und dann selbstständig den Prozess fortführen.

Das Ziel, durch die KI den Zuckerwert zu bestimmen und als Prozessleitgröße zu verwenden, konnte größtenteils

umgesetzt werden. Zusätzlich können anhand der Messwerte die Vorgänge und Rezepte im bIS optimiert und angepasst werden. Beispielsweise kann im Rezept anstatt einer Dauer ein Zielwert für den Zucker angegeben und somit der Prozess flexibler gestaltet werden. Für den Bierbrauer können damit ein zu langes Heizen und damit verbundene Kosten vermieden werden.

Eine große Herausforderung in jedem KI-Projekt ist das Finden einer passenden Netzarchitektur und die passende Vorbereitung der Daten. Dazu wurden während des Projekts verschiedene Ansätze überprüft. Wobei sich grundsätzlich kleine Netze als vorteilhaft erwiesen haben, da durch das schnelle Training viele Kombinationen aus Hyperparametern und Netzarchitekturen getestet werden konnten und dies bei Ergebnissen, welche denen großer Netze ähnlich sind.

Das finale Netz liefert teilweise sehr gute Ergebnisse, teilweise aber auch noch verbesserungswürdige. Dabei ist oftmals ein Unterschied in den Eingangsdaten nur schwer erkennbar. Die Stabilität und Genauigkeit der Vorhersagen können durch das weitere Sammeln und Trainieren von Daten verbessert werden.

Acknowledgement

Dieses Forschungsprojekt wird von verschiedenen Partnern unterstützt. Die Firma 42aaS GmbH unterstützt in Sachen KI und Infrastruktur mit ihrem Wissen und Arbeitskraft. Die Firma top flow stellt ihr Produkt zur Betriebsdatenerfassung zur Verfügung und hilft dieses an entstehende Anforderungen anzupassen. Zusätzlich hat die Firma Speidel ihr Brau-Equipment und ihre Erfahrung im Brauprozess zur Verfügung gestellt. Die Firma ifm hat verschiedene Sensoren bereitgestellt.

Dieses Projekt wurde vom Ministerium für Wirtschaft, Arbeit und Tourismus des Landes Baden-Württemberg im Rahmen ihres Programms Initiative Wirtschaft 4.0 BW gefördert und finanziert

LITERATUR

- [1] Otto Föllinger
Regelungstechnik, Einführung in die Methoden und ihre Anwendungen
Erschienen: 01.06.2008
Zugriff: 20.09.2021
<https://www.vde-verlag.de/buecher/604201/regelungstechnik.html>
- [2] D. G. Luenberger
Observing the state of a linear system.
Erschienen: 01.04.1964
Zugriff: 29.09.2021
IEEE Transaction on Military Electronics. Band 8
<https://ieeexplore.ieee.org/document/4323124>
- [3] Brian Gazaille, Benjamin Bayer
Soft Sensors for Bioprocess Monitoring
Erschienen: 28.05.2020
Zugriff: 21.10.2021

<https://bioprocessintl.com/analytical/pat/soft-sensors-for-bioprocess-monitoring/>

- [4] Wolfgang Samenfink, Andreas Kufferath
Verfahren zur Bestimmung der Ölqualität
Erschienen: 07.12.2007
Zugriff: 05.09.2021
<https://patents.google.com/patent/DE102006059071A1/de>

- [5] Fred Clavel
Basics: Standardization and the Z score
Erschienen: 18.03.2019
Zugriff: 18.11.2021
<https://fredclavel.org/2019/03/18/basics-standardization-and-the-z-score/>

KONTAKT

Malte Hoffmann, BS, studierte an der Hochschule Albstadt-Sigmaringen Informatik und arbeitete für ein Forschungsjahr am Projekt KI-VISOPRO.

Derk Rembold ist Professor an der Hochschule Albstadt-Sigmaringen und unterrichtet technische Informatik und forscht mit Schwerpunkt Cyber Physical Systems.

Bernd Stauß ist Professor an der Hochschule Albstadt-Sigmaringen und unterrichtet Wirtschaftsinformatik. Sein Forschungsschwerpunkt liegt im Bereich Business Analytics.

Stefan Schwarzkopf ist Geschäftsführer der Firma 42aaS GmbH und beschäftigt sich insbesondere mit KI Themen bei seinen Kundenprojekten.

Marius Jacobi studierte an der Hochschule Albstadt-Sigmaringen technische Informatik und bearbeitet seine BS Thesis bei der Firma 42aaS mit dem Thema KI.

Robotic Process Automation als Enabler der Digitalen Transformation der OSRAM Continental GmbH

B.A. Elena Ruf¹
Oliver Dittert¹
Professor Dr. Frank Herrmann²

¹ OSRAM Continental GmbH
Purchasing Operations
Im Gewerbepark 25C
93059 Regensburg
E-Mail: ruf-elena@outlook.de

² Ostbayerische Technische Hochschule Regensburg (OTH Regensburg)
Innovationszentrum für Produktionslogistik und Fabrikplanung
Postfach 120327, 93025 Regensburg
E-Mail: Frank.Herrmann@OTH-Regensburg.de

SCHLÜSSELWÖRTER

Robotic Process Automation (RPA), Business Process Automation (BPA), Digitale Transformation

ABSTRACT

Der vorliegende Artikel befasst sich mit der Automatisierung von Unternehmensprozessen mit Hilfe von Robotic Process Automation (RPA). Um die zentrale Frage nach den Potentialen von RPA zu beantworten, wurde gemeinsam mit der OSRAM Continental GmbH eine Fallstudie in der Einkaufsabteilung durchgeführt. In dieser Fallstudie werden die Fähigkeiten von Attended RPA sowie die Unterschiede im Vergleich zu einem bereits im Unternehmen bestehenden Workflow-Automation-Systems untersucht. Die Automatisierung hat gezeigt, dass mittels RPA das Automatisierungspotential eindeutig erweitert werden kann. Jedoch wurden ebenso deutlich, dass die Technologie eindeutig auf bestimmte Aufgaben begrenzt ist und daher ein hybrider Automatisierungsansatz mit beiden Technologien die beste Lösung darstellt.

EINLEITUNG

Sowohl die Wirtschaft als auch unsere Gesellschaft befinden sich im Wandel: Getrieben durch den technologischen Fortschritt durchdringen neue, innovative und disruptive Technologien wie beispielsweise Künstliche Intelligenz (KI), Internet of Things, Cloud Computing, Big Data oder Blockchain unseren Alltag und verändern unser Leben sowie die Arbeitsweisen von Unternehmen (Oswald und Krcmar 2018).

Aufgrund der schnellen Entwicklung neuer und bereits bestehender Technologien, müssen Unternehmen bereit sein, ihre Arbeitsweisen zu überdenken, um die Potentiale der Digitalisierung nutzen und somit nachhaltiges Wachstum sicherstellen zu können. Diese Anpassung an neue Technologien wird auch als *Digitale Transformation* bezeichnet und wird häufig mit den Schlagwörtern „Unausweichlichkeit, Unumkehrbarkeit, ungeheure

Schnelligkeit und Unsicherheit in der Ausführung“ charakterisiert (Oswald und Krcmar 2018, Gimpel und Röglinger 2015). Es wird daher deutlich, dass es sich um einen ständigen Transformationsprozess handelt, an dem Unternehmen aktiv teilnehmen müssen, um in der digitalen Welt zu bestehen und von deren Vorteilen zu profitieren (zukunftsInstitut 2021). Dabei stellt die Automatisierung und Digitalisierung von Prozessen eine zentrale Dimension der Digitalen Transformation dar (Reinhardt 2020).

In diesem Kontext rückt insbesondere *Robotic Process Automation* immer mehr in den Fokus.

In Literatur und Praxis wird von großem Wachstum und den Erfolgen im Zusammenhang mit der Technologie berichtet: So beschreibt Gartner den RPA Markt als „eines der schnellst wachsenden Segmente im Softwarebereich“. Und schätzt, dass „der Markt aufgrund seines zweistelligen Wachstums erwartungsgemäß bis 2024 einen Wert von über drei Milliarden USD erreichen wird“ (Rashid und Ray 2021). Das Ziel des Einsatzes der Software-Robots ist es, den Menschen von manuellen und repetitiven Tätigkeiten zu befreien und somit mehr Freiraum für Kreativität, Interaktion und Interpretation zu schaffen. Weiterhin wird in der Literatur im Zusammenhang mit RPA von einem hohen Return on Investment (ROI) berichtet. Bereits im ersten Jahr sollen bis zu 200% je nach Prozess möglich sein (Willcocks et al. 2019). Aufgrund dieser Faktoren kam PwC im Rahmen einer RPA Studie in der DACH-Region im Jahr 2020 zu folgendem Entschluss: „Um langfristig wettbewerbsfähig zu bleiben und um nachhaltig zu wachsen, sollten Unternehmen sich mit der Automatisierung ihrer Prozesse auseinandersetzen, um Fachkräftemangel entgegenzuwirken und Mitarbeitern mehr Zeit für wertschöpfende, statt transaktionale Tätigkeiten zu schaffen“ (Bardens 2020). RPA ist somit in aller Munde. Gleichzeitig stellt sich jedoch auch die Frage, ob es sich nur um einen kurzfristi-

gen Hype handelt oder, ob wir bald – vergleichbar zu Industrierobotern - einen großflächigen Einzug der virtuellen Software-Robots in unsere Büros erleben werden (Suhr 2020).

Gegenstand dieser Arbeit ist daher die zentrale Frage nach den Potentialen der robotergesteuerten Prozessautomatisierung. Für die Beantwortung dieser Frage wurde in Zusammenarbeit mit der Einkaufsabteilung der OSRAM-Continental GmbH (OC) eine Fallstudie durchgeführt, um die Technologie genauer zu untersuchen und von den bestehenden Automatisierungsmöglichkeiten im Unternehmen abgrenzen zu können.

AUFBAU

Im Folgenden wird zunächst der zu automatisierende Prozess beschrieben und analysiert. Im Anschluss werden die zwei Technologien Workflow-Automation und RPA aus dem Bereich der Business Process Automation vorgestellt, wobei RPA aufgrund von Neuheit und Relevanz im Verlauf genauer betrachtet wird. Die Eignung des Prozesses und die Implementierung der Automatisierung wird erläutert und abschließend werden die Ergebnisse und allgemeinen Erkenntnisse präsentiert.

LEAD-TIME-PROZESS: BESCHREIBUNG UND ANALYSE

OC verwendet zum aktuellen Zeitpunkt bereits das ERP-System SAP für die einfachere und effizientere Verarbeitung von Daten und die Organisation der wesentlichen Unternehmensfunktionen. Durch Customizing konnte außerdem für viele Prozesse ein hoher Automatisierungsgrad erreicht werden. Weiterhin wurde vor Kurzem das Microsoft Workflow-Automation-Programm „Power Automate“ in der Einkaufsabteilung eingeführt. Jedoch existieren weiterhin Prozesse und Teilschritte, welche die manuelle Durchführung durch einen Mitarbeitenden des Unternehmens erfordern.

Für die Untersuchung und Automatisierung im Rahmen dieser Arbeit wurde der Lead-Times (LT) Prozess aus dem Werkseinkauf ausgewählt, welcher zum einen als sehr zeitintensiv eingeschätzt wurde. Zum anderen erschien der Prozess aufgrund der verwendeten Applikationen und Tätigkeiten besonders interessant für die Übertragung der Untersuchungsergebnisse auf zukünftige Automatisierungskandidaten.

Die Abfrage der LTs liegt bei OC in der Verantwortung der Werkseinkäufer der Produktivstandorte und ist für verschiedene Tätigkeiten des Supply Chain Managements entscheidend. Als grundlegender Parameter nehmen die LT Einfluss auf die Planung von Aufträgen und Produktion, sowie in Folge die Abgabe von Bestellungen bei den Lieferanten. Bei diesen sogenannten LT handelt es sich um eine variable Zeitangabe, welche die Dauer zwischen der Bestellung bei dem Lieferanten und dem Eintreffen der Komponenten am Standort beschreibt. Diese Zeitangabe ist jedoch eine instabile Information, welche von verschiedenen Faktoren insbesondere auf Seiten des Lieferanten beeinflusst wird. Aufgrund dieser Abhängigkeit können die LT nur bedingt durch die Ein-

kaufsabteilung beeinflusst werden. Stattdessen ist es entscheidend für eine erfolgreiche Produktion, dass die LT in regelmäßigen Abständen abgefragt werden und dass die aktuellen Daten ständig in den internen Systemen abgerufen werden können.

Um diese Verfügbarkeit gewährleisten zu können, werden die LT regelmäßig von den Produktivstandorten angefragt, geprüft und schließlich in SAP aktualisiert. Der in Abbildung 1 dargestellte Prozessablauf zeigt die erforderlichen Arbeitsschritte und ist an den betroffenen Standorten vergleichbar, jedoch finden sich kleinere Unterschiede aufgrund verschiedener SAP-Systeme.

Wie Abbildung 1 zu entnehmen ist, werden zu Beginn des Monats zunächst alle relevanten Daten über einen Massendownload aus SAP extrahiert. Diese Daten werden daraufhin in eine Reporting-Datei übertragen und in einzelne Übersichten aufgeteilt, sodass anschließend eine Anfrage an jeden Lieferanten mit aktiven Komponenten gestellt werden kann. Nach Angaben einer Mitarbeiterin, welche für die Ausführung des LT-Prozesses am Standort Treviso verantwortlich ist, werden für die Vorbereitung und das Senden aller Anfragen an insgesamt 67 Lieferanten etwa 1,5 bis 2 Arbeitstage benötigt. Gemäß der Erfahrung der Mitarbeitenden erhält OC von einigen Lieferanten zunächst keine Rückmeldung oder unvollständige Daten und muss daher nochmals Rücksprache mit diesen Lieferanten halten. Weiterhin geben Lieferanten die LT häufig in einem falschen Format an, weshalb eine zusätzliche Überarbeitung der Daten notwendig sein kann. Sobald die Daten jedoch überprüft und gegebenenfalls angepasst wurden, werden sie so schnell wie möglich formatiert und schließlich im SAP-System hochgeladen. Insgesamt läuft das Tracking der Rückmeldungen, Senden von Erinnerungen sowie die benötigten Datenqualitätschecks, Anpassungen und Uploads in das SAP-System über einen Zeitraum von ca. 2 Wochen.

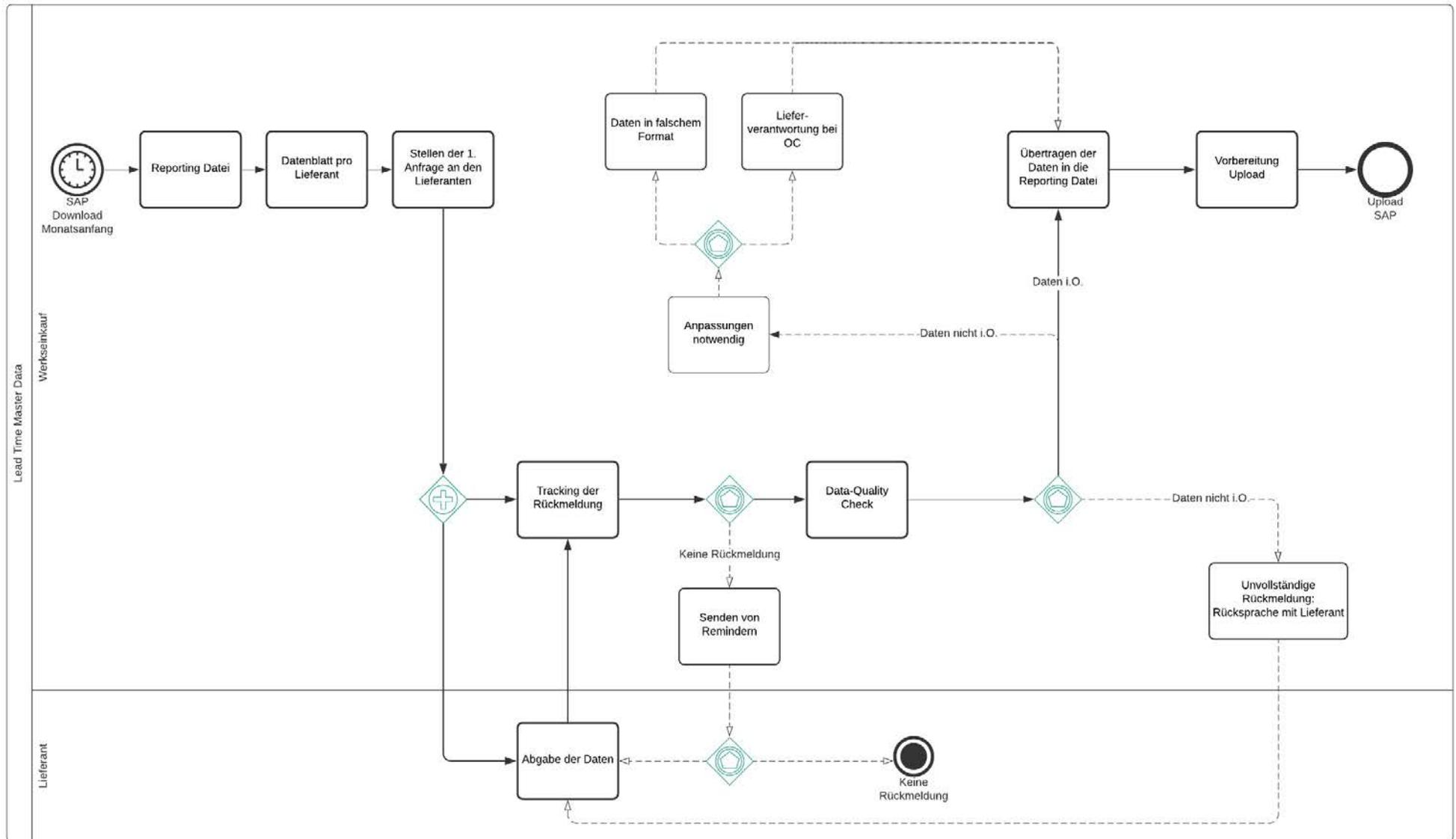


Abbildung 1 Lead Times Prozess vor der Automatisierung

Über eine detaillierte Analyse der Einzelschritte konnte schließlich die Gesamtdauer des Prozesses abhängig von der Anzahl der Lieferanten ermittelt werden. Am Standort Treviso wird bei der Ausführung nach Idealfall für insgesamt 67 Lieferanten von einer Gesamtdauer von 26-39 Stunden pro Monat ausgegangen. Für die Berechnung der Kosten wird mit 2 Szenarien gearbeitet: Der Prozess kann an einem Best Cost oder einem High Cost Standort durchgeführt werden, wobei im Folgenden Pauschalbeträge für die Berechnung der Prozesskosten verwendet werden. Bei der Durchführung an einem Best Cost Standort werden Prozesskosten von ca. 860-1.290€ pro Monat angenommen. Da der Prozess aktuell von Mitarbeitenden an High Cost Standorten durchgeführt wird, werden die monatlichen Kosten auf ca. 1.690-2.535€ geschätzt. Jedoch müssen bei diesen Angaben die nachfolgenden Aspekte berücksichtigt werden.

Es wird angenommen, dass der Prozess in Realität eine etwas höhere Gesamtdauer aufweist, als hier berechnet wurde. Zum einen enthält der Prozess viele repetitive Einzelschritte, bei welchen sich häufige Leichtsinnsfehler einschleichen können und Verbesserungen oder Datenqualitätschecks notwendig werden - was sich wiederum negativ auf die Gesamtdurchlaufzeit des Prozesses auswirkt. Zum anderen wird gemäß Angaben der verantwortlichen Mitarbeitenden ein Großteil des Prozesses ad hoc oder simultan zu anderen Tätigkeiten ausgeführt, da die Rückmeldungen der Lieferanten zu unterschiedlichen Zeitpunkten bei OC eingehen, jedoch so schnell wie möglich verarbeitet werden sollen. Multitasking führt in der Regel allerdings zu einer geringeren Konzentrationsfähigkeit, Zeitverlust und einer insgesamt geringeren Qualität der Arbeit, was den Gesamtaufwand des Prozesses erhöht. Weiterhin ist zu berücksichtigen, dass die Ausführung des LT Prozesses somit gleichzeitig auch den Zeitaufwand und die Qualität anderer, durch das Multitasking beeinträchtigte Tätigkeiten negativ beeinflussen könnte (Moritz und Rimbach 2006).

Außerdem sollte auch die Realität in den Werken bei der Analyse des Prozesses berücksichtigt werden. In der aktuellen Situation, aufgrund der Corona-Pandemie, sollten die LTs jeden Monat bei allen Lieferanten angefragt werden. Jedoch fehlen Kapazitäten, was dazu führt, dass der Prozess nicht für alle Lieferanten jeden Monat vollständig durchgeführt werden kann. Aufgrund sich schnell verändernder LT bedeutet dies jedoch, dass womöglich Bestellungen zum falschen Zeitpunkt abgegeben werden, wodurch es im schlimmsten Fall zu Bandstillständen in der Produktion kommen könnte. Diese Opportunitätskosten lassen sich jedoch nicht eindeutig quantifizieren.

Um Mitarbeitende zu entlasten und das Abfragen der LT in regelmäßigen Zeitabständen sicherzustellen, sollen die manuellen Tätigkeiten des Prozesses im Rahmen dieser Arbeit automatisiert werden.

Diese Automatisierung soll mit Hilfe des bereits im Unternehmen verbreiteten Workflow-Automation-System und einer Attended RPA Software vorgenommen werden. Die beiden Technologien aus dem Bereich der Business Process Automation werden im Folgenden vorge-

stellt. Die Hintergründe und Anforderungen für den Einsatz von RPA werden aufgrund der Neuheit und Relevanz der Technologie näher betrachtet.

AUTOMATISIERUNGSANSÄTZE

Business Process Automation

Bei Business Process Automation (BPA) handelt es sich um einen Oberbegriff für den Einsatz verschiedener Technologien und Methoden zur Automatisierung von Prozessen in einem Unternehmen. Dabei liegt der Fokus auf der effizienten und schnellen Durchführung insbesondere administrativer Prozesse (Koch und Fedtke 2020). Außerdem soll durch den Einsatz der Technologien eine Qualitätssteigerung erreicht werden und der Prozess somit insgesamt optimiert werden (Mohapatra 2013).

Workflow Automation

Eine bereits länger verbreitete Methode aus dem Bereich der BPA sind Workflow-Automation-Systeme. Ein Workflow bezeichnet zunächst unterschiedliche, eigenständige Handlungen, welche in einer bestimmten Reihenfolge ausgeführt werden, um einen Unternehmensprozess durchzuführen (Georgakopoulos et al. 1995). Diese einzelnen Elemente des Workflows werden dabei entweder von Mitarbeitenden eines Unternehmens, Software-Systemen oder in einer Kombination bearbeitet. Für die Workflow-Automation werden die in den einzelnen Arbeitsschritten verwendeten Software-Systeme über Schnittstellen (APIs) in ein Workflow-Automation-System eingebunden und können somit im Ausführungsverlauf automatisch gestartet werden und miteinander kommunizieren (Georgakopoulos et al. 1995).

Bei dem im Unternehmen verwendeten Workflow-Automation-System handelt es sich um die Microsoft Software Power Automate. Microsoft selbst bezeichnet Power Automate als „Cloud-based digital process automation (DPA)“ Tool (Microsoft 2020). Alternativ lässt sich Power Automate auch als Web-Automatisierung beschreiben, welche auf der Verbindung von Applikation über technische Schnittstellen (APIs) basiert. Daher ist zu beachten, dass hauptsächlich moderne Web-Applikationen mit einer API in der Automatisierung integriert werden können. Microsoft gibt dabei an, dass bereits auf über 380 Applikationen zugegriffen werden kann und es besteht zudem die Möglichkeit, selbst weitere Konnektoren zu entwickeln (Microsoft 2020). Außerdem handelt es sich bei Power Automate um eine Low-Code/No-Code Plattform, weshalb kaum bis keine Programmierkenntnisse zum Erstellen der Workflows notwendig sind (Microsoft 2020).

Robotic Process Automation (RPA)

Robotic Process Automation (RPA), zu Deutsch roboter-gesteuerte Prozessautomatisierung, ist ebenfalls eine Technologie aus dem Bereich der Business Process Automation (BPA), welche die Automatisierung strukturierter, digitaler Arbeitsabläufe ermöglicht, ohne Veränderungen an bestehenden Anwendungen vorzunehmen (Koch und Fedtke 2020, Langmann und Turi 2020).

Die Automatisierung basiert dabei jedoch nicht, wie die Bezeichnung fälschlicherweise vermuten lässt, auf klassischen Robotern, welche man aus der Fertigung kennt, sondern auf virtuellen Software-Robotern. Diese sogenannten „Bots“ werden im Rahmen einer Automatisierung darauf trainiert, die Aktionen von Menschen in verschiedenen Software-Applikationen nachzuahmen (Lacity und Willcocks 2016). Die Roboter agieren dabei weiterhin nur über die Benutzeroberfläche – auch Graphical User Interface (GUI) genannt – genauso wie Mitarbeitende, welche die Prozesse zuvor ausgeführt haben (Van der Aalst et al. 2018, Plattfaut et al. 2020). Durch dieses Vorgehen werden keine Änderungen an den zugrundeliegenden Strukturen im Backend-System vorgenommen, weshalb RPA auch als Lightweight IT kategorisiert wird (Bygstad 2017).

Im Vergleich zu klassischen Automatisierungen können mit dieser GUI Automatisierung zeitaufwändige Systemintegrationen oder eine Verbindung der Systeme mittels technischer Schnittstellen (APIs) umgangen werden, wodurch eine schnellere Implementierung ermöglicht wird. Allerdings ist gleichzeitig zu beachten, dass RPA durch dieses Vorgehen auf zugrundeliegenden Systemen aufbaut und somit auch von diesen abhängig ist (Koch und Fedtke 2020, Langmann und Turi 2020).

RPA gilt weiterhin als Low-Code/No-Code Lösung, da für die Implementierung der Bots üblicherweise kaum bis keine Programmierkenntnisse erforderlich sind. Die RPA-Plattformen arbeiten stattdessen mit bereits vorgefertigten Drag-and-Drop Bausteinen, welche die einzelnen Arbeitsschritte modellieren und von Nutzern an den individuellen Use Case angepasst werden können. Häufig wird außerdem - ähnlich wie bei Excel Macros - eine Aufnahmefunktion zur Verfügung gestellt, welche das Zusammenstellen des Roboters ebenfalls erleichtern und beschleunigen kann (Taulli 2020).

Zusammenfassend ist RPA somit eine schnelle und einfach umzusetzende Möglichkeit der Automatisierung (Penttinen et al. 2018). Jedoch ist zu beachten, dass nicht alle Prozesse für diese Art der Prozessautomatisierung geeignet sind. Denn die Roboter sind beispielsweise nicht fähig, eigenständig eine Logik oder eine Bedeutung aus Daten abzuleiten und sind daher hauptsächlich für strukturierte und regelbasierte Routinetätigkeiten konzipiert. Um anspruchsvollere Aufgaben zu automatisieren, können jedoch zusätzliche Technologien oder Künstliche Intelligenz angewendet werden (Sikora et al. 2019, Penttinen et al. 2018).

RPA FÜR DIE PROZESSAUTOMATISIERUNG

Stand der Entwicklung

Es existieren bereits die unterschiedlichen Automatisierungsvarianten Attended RPA, Unattended RPA und Intelligent Process Automation (IPA) oder Cognitive Automation. Für die erfolgreiche Implementierung ist es wichtig, die Unterschiede zwischen den Varianten zu kennen (Taulli 2020).

Attended RPA

Attended RPA oder Robotic Desktop Automation (RDA) bezeichnet eine Form von RPA, welche lokal auf dem Gerät bzw. auf dem Desktop des Anwenders ausgeführt wird. Bei dieser Variante agiert der Roboter wie ein virtueller, persönlicher Assistent des Anwenders. Der Benutzer kann den Roboter nach seinen individuellen Bedürfnissen gestalten, ihn eigenständig starten und beispielsweise Arbeitsschritte automatisieren, welche eine Interaktion zwischen Menschen und Roboter erfordern. Da der Software-Roboter jedoch lokal auf dem Gerät eingesetzt wird und auch lokal die Steuerung der Applikationen übernimmt, kann der Anwender nicht oder nur eingeschränkt weiterarbeiten, während der Roboter seine Aufgaben ausführt. Gleichzeitig entsteht durch die lokale Verwendung ein erhöhter Schulungsaufwand und das zentrale Tracking des Erfolges, der Vorteile und der Risiken der Maßnahme wird erschwert (Koch und Fedtke 2020, Langmann und Turi 2020, Taulli 2020).

Unattended RPA

Wie der Name bereits verdeutlicht, erfolgt die Automatisierung mit Unattended RPA ohne direkte Beaufsichtigung oder Interaktion mit einem Anwender. Die Bots starten ihre Tätigkeit automatisch, sobald sie durch einen im Voraus festgelegten Trigger aktiviert werden und arbeiten dabei zentral über einen Server auf virtuellen Maschinen. Bei dieser Form der RPA besitzt der Roboter in der Computerumgebung eine eigene Identität und agiert somit als virtueller Mitarbeiter, welcher seine Tätigkeit eigenständig startet und ausführt. Auch die Entwicklung und Verwaltung des Roboters liegt bei dieser RPA Variante für gewöhnlich nicht mehr in der Hand der einzelnen Anwender, sondern wird von einem für RPA verantwortlichen Team übernommen. Ein Nachteil dieser Form liegt daher darin, dass sich die RPA-Verantwortlichen immer wieder in neue Prozesse hineinarbeiten müssen, um diese erfolgreich automatisieren zu können. Gleichzeitig ist jedoch die Überwachung des Erfolges und der Qualität der Maßnahme durch eine zentrale Verwaltung deutlich effizienter und akkurater. Weiterhin wird bei der zentralen Verwaltung durch ein Experten-Team auch das Wissen zentral gespeichert, wodurch nicht nur eine schnellere und qualitativ hochwertigere Entwicklung der Roboter möglich ist, sondern auch die Weiterverwendung bereits existierender Roboter (Koch und Fedtke 2020, Langmann und Turi 2020, Taulli 2020).

Intelligent Process Automation

Intelligent Process Automation (IPA)/ Cognitive RPA Sowohl Attended als auch Unattended RPA sind nicht dazu fähig, eigenständig komplexe Entscheidungen zu treffen oder eine Logik aus Daten abzuleiten. Daher ist der Anteil an geeigneten Prozessen für die Automatisierung mit diesen RPA Varianten begrenzt. Um jedoch diese Schwachstellen zu überwinden und auch komplexere Prozesse automatisieren zu können, wird RPA mit Künstlicher Intelligenz (KI) verbunden. In diesem Zusammenhang spricht man von Intelligent Process Automation (IPA) oder Cognitive RPA. Ein besonderer Fokus

für den Einsatz von IPA liegt auf den zu verarbeitenden Daten. Attended sowie Unattended RPA können derzeit fast ausschließlich mit strukturierten Daten arbeiten. In den meisten Unternehmen ist jedoch der Anteil an unstrukturierten Daten deutlich höher. Die Integration von KI wird daher als Chance gesehen, um das RPA-Automatisierungspotential zu erweitern, da mit Hilfe von Technologien wie NLP, Sprach oder Bilderkennung auch Prozesse mit unstrukturierten Daten automatisiert werden können. Gleichzeitig ist jedoch zu beachten, dass mit dem Einsatz weiterer Technologien nicht nur der Handlungsspielraum, sondern auch die Komplexität der Automatisierung zunimmt (Koch und Fedtke 2020, Langmann und Turi 2020, Czarnecki und Auth 2018).

Vorteile RPA

Die unterschiedlichen Varianten von RPA bringen je nach Use Case auch unterschiedliche Vor- und Nachteile mit sich. Da für OC zunächst besonders die Varianten Attended und Unattended RPA von Interesse sind, werden im Folgenden die grundlegenden Vorteile der beiden Varianten erläutert.

Zeitersparnis und Produktivitätssteigerung

Der wohl offensichtlichste Vorteil und Grundgedanke hinter einer Automatisierung ist die Zeitersparnis, welche mit Hilfe der Maßnahme erreicht werden soll. Beim Einsatz von RPA übernehmen virtuelle Roboter zeitaufwändige, häufig anfallende aber geistig weniger anspruchsvolle Tätigkeiten. Somit werden Mitarbeitende entlastet und können sich den tatsächlich wertschöpfenden Aufgaben ihrer Funktion widmen. Demzufolge ist bei einer robotergesteuerten Prozessautomatisierung nicht nur mit Zeitersparnis zu rechnen, sondern auch mit einer daraus resultierenden Produktivitätssteigerung, da Mitarbeitende mehr Zeit in kreative, interaktions- oder interpretationsbasierte Aufgaben investieren können (Bardens 2020, Kroll et al. 2016).

Qualität

Gleichzeitig wird im Zuge einer RPA häufig auch eine Qualitätssteigerung bei den automatisierten Prozessen beobachtet. Die für RPA geeigneten Prozesse zeichnen sich durch einen repetitiven Charakter mit vielen einzelnen Teilschritten aus. Bei der Ausführung durch den Menschen schleichen sich hier oftmals Leichtsinnsfehler ein, wie beispielsweise Klick und Tipp-Fehler oder das Überspringen von Teilaufgaben, was eine häufige Nachbesserung und Qualitätschecks zur Folge hat. Der Roboter hält sich im Gegensatz bei jeder einzelnen Ausführung genau an zuvor definierte Regeln und Abfolgen der Aufgaben. Er hat daher bei richtigen Dateninputs eine weitaus geringere Fehlerquote und verhält sich gleichzeitig stets regelkonform, wodurch insgesamt die Qualität des Prozesses gesteigert werden kann (Bardens 2020, Kroll et al. 2016).

Verfügbarkeit und Skalierbarkeit

Der Roboter wird außerdem im Vergleich zum Menschen nicht müde und ist, sofern es keine technischen Schwierigkeiten gibt, jeden Tag 24 Stunden einsatzbereit – 365 Tage im Jahr. Gleichzeitig ist die einfache Skalierbarkeit ein weiterer großer Vorteil der Maßnahme. Sollten die Roboter trotz ständiger Verfügbarkeit den Bedarf nicht decken können, so können die Kapazitäten durch die Implementierung weiterer Roboter aufgestockt werden (Tauli 2020, Bardens 2020, Kroll et al. 2016).

Wirtschaftlichkeit

Einer der wichtigsten Vorteile von RPA, auch im Vergleich zu anderen Automatisierungsmöglichkeiten, ist die Wirtschaftlichkeit der Maßnahme. RPA greift bei der Automatisierung nicht in die tiefgehenden Strukturen im Backend-System ein, sondern agiert nur über das GUI (Bygstad 2017). RPA-Plattformen ermöglichen daher eine schnelle und einfache Entwicklung der Roboter, welche im weiteren Verlauf außerdem leicht an wechselnde Bedürfnisse angepasst werden kann. Das Zusammenspiel aus Schnelligkeit und geringer Komplexität macht RPA somit zu einer kostengünstigen und gleichzeitig agilen Automatisierungslösung (Penttinen 2018). In Literatur und Praxis wird daher häufig von einem hohen Return on Investment (ROI) berichtet, was die Attraktivität von RPA für Unternehmen steigert (Langmann und Turi 2020, Van der Aalst 2018).

Eignung eines Prozesses

In der Literatur findet sich für die Entscheidungsfindung und Auswahl geeigneter Prozesse bereits eine große Anzahl an Kriterien. Im Folgenden wurden insgesamt 15 Veröffentlichungen analysiert und die am häufigsten genannten Faktoren herausgearbeitet.

Grundlegende Voraussetzungen

Bei der Auswahl der Prozesse für Attended/Unattended RPA sollten die folgenden Kriterien unbedingt erfüllt sein, um die Umsetzbarkeit der Automatisierung gewährleisten zu können.

Wenn-Dann-Faktor

Damit der Roboter den Prozess später reibungslos durchlaufen kann, muss bei der Entwicklung für jede mögliche Handlungsalternative Schritt für Schritt jede durchzuführende Aktion dokumentiert werden. Daher sollten nur regelbasierte Prozesse für die Automatisierung ausgewählt werden, welche für alle möglichen Szenarien eine klare Abfolge von Handlungen nach dem Wenn-Dann-Prinzip vorgeben (Langmann und Turi 2020).

Ausgereifte Prozesse

Für eine möglichst schnelle und erfolgreiche Implementierung ist es außerdem entscheidend, dass die einzelnen Arbeitsschritte bereits im Voraus erprobt beziehungsweise optimiert wurden und keine unbekanntes Szenarien auftreten. Es sollten daher ausgereifte Prozesse ausgewählt werden, welche bereits standardisiert sind. Auf diese Weise kann gewährleistet werden, dass alle Ar-

beitsschritte im Prozess genau bekannt und gut dokumentiert sind. Gleichzeitig sinkt die Wahrscheinlichkeit, dass kurz nach der Implementierung Änderungen am Prozess notwendig werden, was die Stabilität des Roboters gefährden könnte und womöglich eine erneute Überarbeitung des Roboters zur Folge hätte (Langmann und Turi 2020).

Daten

Die genaue Untersuchung der im Prozess zu verarbeitenden Daten ist wichtig bei der Auswahl geeigneter Automatisierungskandidaten. Wie bereits erläutert, können die RPA Varianten ohne KI-Technologien nicht mit unstrukturierten Daten arbeiten. Soll auf den Einsatz von IPA verzichtet werden, können im Prozess ausschließlich strukturierte Daten, welche immer in einem bekannten, vordefinierten Format vorliegen, verarbeitet werden (Plattfaut et al. 2020, Langmann und Turi 2020). Weiterhin müssen alle relevanten Daten in digitaler Form verfügbar sein, da der virtuelle Roboter ansonsten ohne manuelles Eingreifen keinen Zugriff auf die Daten außerhalb der Computerumgebung hat (Plattfaut et al. 2020). Letztlich spielt außerdem die Datenqualität eine große Rolle, denn bei schlechter Datenqualität kann es zu Fehlern bei der Ausführung durch den Roboter kommen (Syed et al. 2020).

Zusatzkriterien für die Implementierung

Während die grundlegenden Voraussetzungen unbedingt erfüllt sein sollten, um die Umsetzbarkeit der Automatisierung zu gewährleisten, sind folgende Faktoren nicht grundsätzlich notwendig. Vielmehr können diese Kriterien von Vorteil für den jeweiligen Use-Case sein und sollten nicht außer Acht gelassen werden, da sie eine schnelle Implementierung und den Erfolg der Maßnahme begünstigen (Langmann und Turi 2020).

Komplexität

Roboter können auch komplexere Prozesse (teil)automatisieren. Gleichzeitig steigt mit der Komplexität jedoch auch der Implementierungs-, Test- und Wartungsaufwand des Roboters (Langmann und Turi 2020). Daher sollten gemäß Gartner (Rashid und Ray 2021) Prozesse für die Automatisierung bevorzugt werden, welche eine geringe Anzahl an Prozessschritten aufweisen. Dabei werden idealerweise weniger als 20 Arbeitsschritte empfohlen (Rashid und Ray 2021). Neben den einzelnen Prozessschritten ist auch die funktionale Logik sowie die Anzahl der Ausführungspfade entscheidend für die Komplexität eines Prozesses. Es sollten daher Prozesse für die Automatisierung bevorzugt werden, die in ihrem Ablauf keine oder nur wenige, aber gleichzeitig konsistente Variationsmöglichkeiten aufweisen (Plattfaut et al. 2020).

Stabilität der Umgebung

Weiterhin empfehlen (Rashid und Ray 2021) und (Plattfaut et al. 2020) Prozesse auszuwählen, an denen maximal 3-4 verschiedene Applikationen beteiligt sind. Die Begrenzung der Applikationen mag zunächst wider-

sprüchlich erscheinen, da beispielsweise die Datenübertragung zwischen vielen verschiedenen Applikationen sehr zeitaufwändig und fehleranfällig ist und eine Automatisierung daher große Effizienzgewinne verspricht. Gleichzeitig gefährden viele Applikationen jedoch die Stabilität des Roboters. Denn je mehr unterschiedliche Applikationen im Prozess integriert sind, desto eher steigt auch die Wahrscheinlichkeit für Veränderungen in der Benutzeroberfläche einer Applikation, beispielsweise durch Updates. Dies wiederum kann zu fehlerhaften Ausführungen oder Abbrüchen und somit zu einem erhöhten Wartungsaufwand des Bots führen (Langmann und Turi 2020).

Anzahl der User

Vergleichbar mit der Anzahl der beteiligten Applikationen kann auch die Anzahl der im Prozess beteiligten User die Stabilität des Roboters gefährden. Zum einen steigt potentiell mit der Anzahl der User auch die Anzahl der verwendeten Applikationen und somit das Risiko für nachträgliche Änderungen. Zum anderen sind mehrere User für gewöhnlich dann beteiligt, wenn sie Informationen beitragen oder Entscheidungen treffen müssen – was ebenfalls die Komplexität des Prozesses erhöht. Es lässt sich daher festhalten, dass Prozesse, an denen nur wenige User beteiligt sind, für die Automatisierung bevorzugt werden sollten (Langmann und Turi 2020).

Anzahl der Ausnahmefälle

Wie bereits näher erläutert, kann der Roboter keine eigenen Entscheidungen treffen und nur das Ausführen, wozu er zuvor trainiert wurde. Eine große Anzahl an Ausnahme- oder Spezialfällen im Prozesse bedeutet daher gleichzeitig, dass während der Implementierung für jedes dieser Szenarien genaue Handlungsanweisung und die Abfolgen der Aktionen hinterlegt und getestet werden müssen. Demnach steigt mit der Zahl der Ausnahmefälle der Implementierungsaufwand und das Risiko, der Handlungsunfähigkeit, wenn nicht alle Szenarien bei der Implementierung erfasst wurden (Langmann und Turi 2020).

Wirtschaftliche Einflussfaktoren

Bei der Investition in RPA treten unterschiedliche Kosten beispielsweise durch Lizenzierung, Beratung, Schulungen oder die Instandhaltung der Roboter auf. Daher sollten zunächst die Prozesse automatisiert werden, welche den größten wirtschaftlichen Effekt erzeugen, sodass möglichst schnell ein positiver ROI erreicht werden kann.

Routinetätigkeiten

Bei der Auswahl der Prozesse sollte daher darauf geachtet werden, häufig auftretende Prozesse zu wählen. Denn je häufiger ein Prozess durchlaufen wird, desto höher sind auch die Erfolge – wie beispielsweise Effizienzgewinne oder Kosteneinsparungen – welche durch eine RPA erzielt werden können (Langmann und Turi 2020).

Prozesskosten

Weiterhin ist für den wirtschaftlichen Effekt der Automatisierung entscheidend, wer den Prozess durchführt. Besonders manuelle Prozesse, die an High Cost Standorten durchgeführt werden, haben großes Kosteneinsparungspotential. Die Prozesskosten werden zudem maßgeblich davon beeinflusst, wie zeitaufwändig ein Prozess ist. Es sollte daher ebenfalls untersucht werden, welchen Anteil der monatlichen Arbeitszeit der Prozess in Anspruch nimmt. Hier gilt erneut die Rechnung: je höher der Anteil, desto höher ist auch das Einsparungspotential durch die Automatisierung, weshalb ein zeitaufwändiger Prozess bevorzugt werden sollte (Plattfaut et al. 2020).

Fehleranfälligkeit

Wie bereits näher erläutert, ist ein großer Vorteil der RPA die Qualitätssteigerung aufgrund der Regelkonformität und der verminderten Fehlerquote. Die Fehleranfälligkeit der Prozesse sollte bei der Auswahl daher ebenfalls analysiert werden. Mit der Übernahme besonders fehleranfälliger Prozesse durch Roboter, können zeitaufwändige Nachbesserungen oder Qualitätschecks vermieden werden, was die Zeit und Kosteneinsparungen durch die Automatisierung erhöht (Osman 2019, Bardens 2020).

Eignung des LT-Prozesses für den Einsatz von RPA

Im Folgenden wird die Eignung des LT-Prozesses für die Automatisierung mit RPA anhand der zuvor definierten Kriterien geprüft.

Grundlegende Voraussetzungen

Der LT-Prozess bietet den Mitarbeitenden nur wenig Freiraum für Kreativität und neue Ideen, da der Prozess immer nach einem genauen Schema ablaufen muss, um erfolgreich durchgeführt werden zu können. Weiterhin hat sich die Prozessausführung schon seit Längerem etabliert und wird an allen Standorten identisch durchgeführt. Der LT-Prozess erfüllt daher die Punkte Wenn-Dann-Faktor und Reife des Prozesses vollständig.

Die Daten stellen eines der Probleme des Prozesses dar – zwar liegen alle verwendeten Daten in digitaler Form vor – jedoch ist die Datenqualität und das Format der Rückmeldungen sehr variabel und wird daher als Risikofaktor für die Implementierung identifiziert. Da die Überprüfung und Anpassung der Daten jedoch bereits ein fester Bestandteil des Prozesses sind, wird die Eignung des Prozesses für eine Automatisierung nicht ausgeschlossen, solange dieser Schritt weiterhin durch einen Mitarbeitenden durchgeführt wird.

Zusatzkriterien für die Implementierung

Der Prozess umfasst bei der Ausführung ohne Ausnahmefälle insgesamt 10 Prozessschritte. Gleichzeitig existieren die 2 Untergliederungen „Keine Rückmeldung durch den Lieferanten“ und „Rückmeldung nicht bereit für den Upload in SAP“, welche zu alternativen Handlungsschritten im Prozess führen. Insgesamt sind der Standardprozess und die Untergliederungen jedoch klar strukturiert und enthalten sich wiederholende Schritte, weshalb der Prozess als wenig komplex eingeschätzt

wird. Im Prozess wird außerdem mit den 3 Applikationen SAP, Excel sowie Outlook gearbeitet, welche sich jedoch sehr selten in ihrer Benutzeroberfläche verändern.

Der Eignungsfaktor Anzahl der User wird ebenfalls als Risikofaktor für die Implementierung eingeschätzt. Insgesamt sind 1-2 OC-Mitarbeitende und alle aktiven Lieferanten am Prozess beteiligt, was direkt mit dem Problem der schlechten Datenqualität der Rückmeldungen zusammenhängt.

Weiterhin existieren mehrere Ausnahmefälle im Prozess, jedoch sind diese bekannt und laufen nach einer bestimmten Vorgabe ab. Da der Prozess allerdings nur zum Großteil und nicht komplett in der Standardausführung abläuft, wird mit einem leicht erhöhten Implementierungsaufwand gerechnet.

Wirtschaftliche Einflussfaktoren

Das Volumen des LT-Prozesses lässt sich nicht eindeutig bewerten. Grundsätzlich wird der Prozess einmal pro Monat durchgeführt. Jedoch sind die einzelnen Arbeitsschritte zeitlich so verteilt, dass während des gesamten Monats sehr häufig am Prozess gearbeitet werden muss. Wie bereits eingangs erläutert, wird der Prozess derzeit von den Werkseinkäufern an den Produktivstandorten durchgeführt. Die Bearbeitungszeit unterscheidet sich an den Standorten aufgrund der unterschiedlichen Anzahl an Lieferanten. Am Beispielstandort Treviso macht der Prozess jedoch bei vollständiger Durchführung nach Idealfall ca. 16-24% der monatlichen Arbeitszeit eines Mitarbeitenden aus. Weiterhin wird der Prozess von den Mitarbeitenden als teilweise fehleranfällig eingeschätzt, da sehr viele Teilschritte beachtet werden müssen. Aus wirtschaftlicher Sicht wird der Prozess daher als sehr geeignet für RPA eingeschätzt.

IMPLEMENTIERUNG DER HYBRIDEN AUTOMATISIERUNGSLÖSUNG

RPA gilt als schnelle und vergleichsweise einfache Möglichkeit der Prozessautomatisierung (Penttinen 2018). Jedoch sind vor Beginn der eigentlichen Umsetzung einige Punkte neben der Auswahl der richtigen Prozesse zu beachten, welche für die Durchführbarkeit und gleichzeitig auch für den Erfolg der Maßnahme entscheidend sind.

Aufgrund des geringen Programmieraufwandes und der einfachen Bedienung wird RPA häufig als Business-Lösung bezeichnet, welche von den betroffenen Abteilungen selbst implementiert werden kann (Czarnecki und Auth 2018, Lacity et al. 2015).

Für die Durchführung der Fallstudie wurde jedoch zunächst die Erlaubnis des Betriebsrates eingeholt, da eine Automatisierung potenziell Arbeitsplätze gefährden könnte. Weiterhin wurden sowohl IT- als auch Rechtsabteilung bezüglich Lizenzierung und Datenschutzvorschriften in die Vorbereitungen der Automatisierung mit einbezogen.

Daraufhin wurde die Automatisierung zunächst geplant und erstellt, getestet und anschließend implementiert. Für die Verwendung der Workflow-Automation mussten im Voraus noch kleinere Änderungen am Prozess vorge-

nommen werden: Da Power Automate Workflow im bestehenden System des Unternehmens weder Daten aus SAP noch Excel auslesen und verarbeiten kann, wurde für den Workflow eine Liste auf der Kommunikationsplattform Microsoft Teams erstellt, welche als Datenbank für das System fungiert. In der Teams Liste kann das Programm die einzelnen Elemente auslesen und die Daten in den Arbeitsschritten verarbeiten, manipulieren und aktualisieren. Der Aufbau und Inhalt der Teams Liste ist dabei an einer bereits bestehenden Tracking-Datei orientiert und enthält alle relevanten Informationen.

Für den Einsatz von RPA mussten keine Veränderungen an den bestehenden Systemen und Vorgehensweisen vorgenommen werden.

Für die Automatisierung des LT-Prozesses wurde schließlich eine hybride Vorgehensweise ausgewählt. Im Prozess kommen daher sowohl Power Automate Workflow als auch Attended RPA zum Einsatz.

Bei der Verwendung dieser Technologien konnte der Prozess jedoch nicht vollständig automatisiert werden und das Eingreifen von Mitarbeitenden ist weiterhin an bestimmten Stellen des Prozesses erforderlich. Die genaue Verteilung der Arbeitsschritte ist Abbildung 2 zu entnehmen.

Der Einsatz der Mitarbeitenden ist im automatisierten Prozess das erste Mal notwendig, sobald die Rückmeldungen der Lieferanten eingehen. Die Überprüfung der Rückmeldungen auf Vollständigkeit und Datenqualität wird weiterhin durch einen Mitarbeitenden aus dem Werkseinkauf durchgeführt. Zum einen ist es wichtig für die Einkäufer, einen Überblick über die LT ihrer Lieferanten zu behalten, zum anderen ist die Rückmeldung der Lieferanten sehr variabel, weshalb ohne die Integration von Technologien der künstlichen Intelligenz die Verarbeitung und Überprüfung sehr aufwändig bis kaum möglich ist, da jeder Spezialfall separat trainiert werden müsste. Aus diesem Grund werden auch Formatanpassungen, bei Auffälligkeiten während des Datenqualitäts-Checks, direkt durch den Mitarbeitenden durchgeführt. Weiterhin nimmt der Mitarbeitende bei spezifischen Rückfragen oder fehlenden Informationen eigenständig zu den betroffenen Lieferanten Kontakt auf. Hier ist menschliche Interpretation der Daten erforderlich und es soll ein persönlicher Kontakt mit dem Lieferanten entstehen. Es wird in diesen Ausnahmefällen keine generische Nachricht erzeugt, um die Beziehung und die gute Zusammenarbeit mit den entsprechenden Lieferanten zu stärken.

Die Wahl des hybriden Automatisierungsansatzes für die restlichen – automatisierbaren – Arbeitsschritte ist vor allem in den Fähigkeiten und Einschränkungen der beiden Technologien begründet.

Mit RPA ist die vollständige Ausführung dieser Arbeitsschritte grundsätzlich möglich, hat sich jedoch bei der Verwendung der Variante Attended RPA als nicht alltagstauglich erwiesen. Bei dem Einsatz dieser RPA Variante muss der Roboter aktiv durch einen Mitarbeitenden gestartet werden und der Mitarbeitende kann während der Ausführung des Roboters nur eingeschränkt oder gar

nicht weiterarbeiten. Power Automate Workflow hingegen wird beispielsweise durch einen zeitlichen Trigger ausgelöst und arbeitet im Hintergrund, ohne Anwendungen zu blockieren. Die vollständige Automatisierung mittels Power Automate Workflow war in der Fallstudie jedoch insbesondere aufgrund fehlender Schnittstellen zu beispielsweise Excel oder den SAP-Systemen nicht möglich. RPA wurde daher immer dort eingesetzt, wo der Einsatz der Workflow-Automation nicht möglich war.

ERGEBNISSE

Mit Hilfe der beiden Automatisierungsmöglichkeiten Power Automate Workflow und Attended RPA konnten insgesamt 11 von 14 Arbeitsschritten auf Seiten von OC automatisiert werden.

Besonders bei den Arbeitsschritten 1-3 konnten mit Hilfe des Roboters große Zeit und Qualitätsverbesserungen festgestellt werden:

Am Standort Treviso und somit bei der Durchführung für 67 Lieferanten konnte der SAP Download, die Datenübertragung in die Reporting-Datei sowie das Erstellen der Datenblätter innerhalb von durchschnittlich 35 Minuten durch den Roboter vorgenommen werden. Bei der Durchführung durch einen Mitarbeitenden ist mit einem durchschnittlichen Zeitaufwand von ca. 250 Minuten ($\approx 4h$) zu rechnen. Die Durchführung dieser Arbeitsschritte mit Hilfe von Attended RPA führt demnach zu einer Zeitersparnis von ca. 85%. Weiterhin ist mit einer Qualitätsverbesserung zu rechnen, da die Dokumente stets nach einem einheitlichen System vorbereitet und in den vorgegebenen Ordnern abgelegt werden.

Auch das Versenden der Anfragen, das Tracking der Rückmeldungen und das Senden von Remindern erfordert mit der Automatisierung nur noch wenige Eingriffe der Mitarbeitenden. Bei Verwendung der hybriden Automatisierungslösung muss das Team lediglich den Workflow zum Update der Liste vor dem 15. Tag jedes Monats starten – die restlichen Schritte erfolgen vollkommen automatisch. Da die Mitarbeitenden jedoch gegebenenfalls den Anfragetext und die Liste bei Aufnahme neuer Lieferanten oder Änderungen bei bestehenden Lieferanten anpassen müssen, kann nicht von einer 99%igen Zeitersparnis ausgegangen werden, was bei Ausführung nach Idealfall zwischen ca. 10-16 Stunden pro Monat entsprechen würde. Der bleibende manuelle Arbeitsaufwand ist jedoch gering und muss in Zukunft näher beobachtet werden, um eine eindeutige Abschätzung treffen zu können.

Bei der Durchführung des Datenqualitätschecks ist – bei der Verwendung von RPA – mit einem erhöhten Arbeitsaufwand zu rechnen. Der Roboter ist sehr sensibel in Bezug auf Formatänderungen und die enthaltenen Daten. Bereits bei einer kleinen Veränderung im Text der Überschrift oder im Format der Rückmeldung kann der Roboter die Daten nicht mehr auslesen und es kommt zu einer Fehlermeldung.

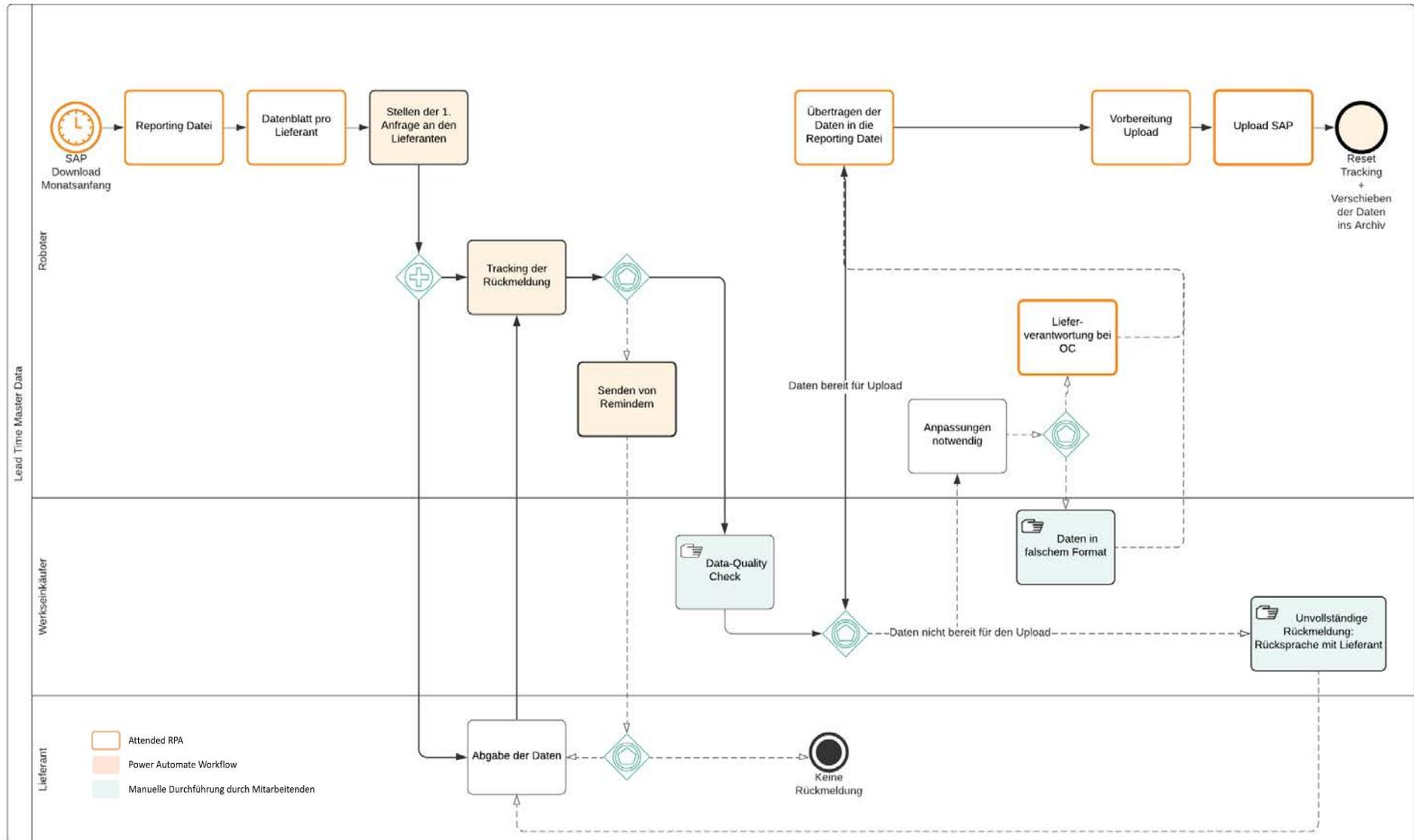


Abbildung 2 Lead Times Prozess Hybride Automatisierungslösung

Die Daten müssen daher von den Mitarbeitern nicht nur inhaltlich, sondern auch in Bezug auf richtige Formatierung geprüft werden. Es wird daher anhand dieser Erkenntnisse angenommen, dass sich der Zeitaufwand für die Überprüfung der Datenqualität um ca. 50% - auf circa 6 Minuten pro Lieferant erhöht.

Wird die Überprüfung der Datenqualität jedoch zuverlässig durchgeführt, so können die folgenden manuellen Arbeitsschritte erneut vollständig durch den Roboter übernommen werden.

Bei der Datenübertragung in Reporting und Upload Datei sowie schließlich bei dem Upload in SAP ist zu beachten, dass die Ausführung durch den Roboter pro Lieferanten nur geringfügig schneller abläuft, da in diesen Schritten hauptsächlich mit UI-Automatisierung gearbeitet wird und der Roboter sich somit, wie ein echter Mitarbeiter, durch die Applikationen klickt und auch auf deren Reaktion warten muss. Aus diesem Grund sollte dem verantwortlichen Mitarbeiter ein separates Endgerät zur Verfügung gestellt oder die Variante Unattended RPA gewählt werden, um die potentiellen Zeitgewinne durch die Übernahme der gesamten Tätigkeit verzeichnen zu können. Jedoch konnte erneut eine Verbesserung der Qualität der Ausführung festgestellt werden, da der Roboter für alle Lieferanten stets alle Arbeitsschritte nach dem zuvor trainierten Schema durchführt und keine Arbeitsschritte auslässt.

Insgesamt konnte mit Hilfe der Automatisierung die Gesamtdurchlaufzeit des Prozesses nur geringfügig verbessert werden. Dies liegt vor allem an der Zerstreuung der einzelnen Aufgaben über einen Zeitraum von ca. drei Wochen und der zeitlich variablen Rückmeldung der einzelnen Lieferanten.

Da jedoch insgesamt 11 von 14 Arbeitsschritten vollständig automatisiert wurden, werden die Mitarbeitenden entlastet und somit die Kosten des Gesamtprozesses gesenkt. Wird ein separates Endgerät oder Unattended RPA zur Verfügung gestellt, so können am Beispielstandort Treviso insgesamt, wenn keine Wartungsarbeiten auftreten, ca. 20-32 Arbeitsstunden pro Monat eingespart werden. Die Kosteneinsparung wird erneut für die beiden Szenarien High Cost sowie Best Cost Standort vor Abzug der RPA-Lizenzkosten berechnet. Erfolgt die Ausführung des Prozesses an einem High Cost Standort, so können durch die Automatisierung unter den genannten Bedingungen pro Monat am Beispielstandort Kosteneinsparungen von 1.300-2.080€ erzielt werden. Bei der Durchführung an einem Best Cost Standort liegen die Einsparungen bei ca. 660-1.060€. Würde der Prozess zum aktuellen Zeitpunkt nach Idealfall durchgeführt werden, so entspricht die Automatisierung einer Kosteneinsparung von durchschnittlich 80% (ohne Einbezug der Software- und Wartungskosten).

Jedoch ist bei dieser Angabe zu beachten, dass zum aktuellen Zeitpunkt die LT nicht an allen Standorten systematisch bei allen Lieferanten angefragt werden. Bei der Implementierung der Automatisierung können somit nicht an allen Standorten Kosteneinsparungen in dem Umfang, wie hier berechnet, erwartet werden.

Jedoch besteht zum aktuellen Zeitpunkt aufgrund der selektiven Abfrage die Gefahr, dass falsche LT verwendet werden, wodurch es im schlimmsten Fall zu Bandstillständen kommen könnte. Durch den Einsatz der Automatisierung kann dieses Risiko vermieden werden und insgesamt eine deutliche Qualitätssteigerung festgestellt werden. Denn die LT werden im Rahmen der Automatisierung zuverlässig und in regelmäßigen Abständen bei allen Lieferanten abgefragt und die Daten können schneller aktualisiert und stets nach einem einheitlichen System verarbeitet und archiviert werden.

VERALLGEMEINERUNG

Die Implementierung der Fallstudie hat insgesamt gezeigt, dass mit Hilfe von RPA das Automatisierungspotential deutlich erweitert werden kann. Es konnte festgestellt werden, dass sich der Roboter an die Bedürfnisse und Vorgehensweisen der Mitarbeitenden anpasst, während bei der Automatisierung mit dem bestehenden Workflow-System der Prozess an die Voraussetzungen und Fähigkeiten des Programms angepasst werden muss. Weiterhin konnte mit RPA jeder Arbeitsschritt automatisiert werden, während dies mit dem Workflow-System nicht möglich war, beispielsweise aufgrund fehlender Schnittstellen.

Besonders bei der Verarbeitung der Rückmeldungen sind jedoch die Nachteile von RPA deutlich geworden. Hier muss stark darauf geachtet werden, dass ein bestimmtes Format eingehalten wird. Kommt es zu Veränderungen, so folgen Fehlermeldungen oder Fehler.

Weiterhin hat die Fallstudie gezeigt, dass Attended RPA eher für kleinere, kürzere Aufgaben geeignet ist oder Aufgaben, welche einen Austausch zwischen Roboter und Mitarbeitenden erfordern. Bei größeren Prozessschritten sollte das Modell Unattended RPA gewählt oder ein separates Endgerät bereitgestellt werden, auf dem der Roboter arbeiten kann, ohne den Mitarbeitenden in seiner Arbeit zu behindern. Andernfalls kann das volle Potential der Automatisierung nicht ausgeschöpft werden.

Weiterhin sind die Position von RPA als Business-Lösung und die Werbeversprechen von Automatisierung ohne Programmierkenntnisse kritisch zu betrachten. Während der Implementierung wurde deutlich, dass die Fachabteilung stets IT- und Rechtsabteilung des Unternehmens in die Automatisierung miteinbeziehen sollte, um Verstöße gegen Lizenzbedingungen und Datenschutz zu vermeiden. Auch in der Literatur wurde dies mittlerweile mehrmals hervorgehoben (Langmann und Turi 2020, Penttinen et al. 2018). Die Fallstudie zeigte außerdem, dass der Roboter erstellt werden kann, ohne selbst Code zu schreiben und auch die Aufzeichnungsfunktion ist sehr hilfreich für das schnelle Zusammenstellen der einzelnen Module. Dennoch werden Grundkenntnisse der Programmierung wie beispielsweise Variablentypen, Listen, Arrays, Schleifen oder Bedingungen für die Implementierung bereits einfacher Automatisierungsaufgaben benötigt.

Ein Werbeversprechen, welches jedoch eindeutig eingehalten werden konnte, ist die kurze Implementierungszeit. Die Automatisierung im Rahmen dieses Projektes

konnte mit Vorkenntnissen des Prozesses, jedoch ohne Vorkenntnisse der RPA-Software innerhalb von ca. 8 Wochen bei einer wöchentlichen Arbeitszeit von ca. 25 Stunden erstellt und getestet werden. Daraufhin wurde die Automatisierung an 2 Standorten implementiert. Während der ersten Ausführung mit Automatisierung wurden die Ergebnisse von der Autorin noch im Detail überprüft und kleinere standortabhängige Anpassungen vorgenommen. Doch bereits ab der zweiten Ausführung konnte die Automatisierung ohne Probleme eingesetzt werden. Es wird davon ausgegangen, dass im Laufe der Zeit Veränderungen notwendig werden, der Wartungsaufwand wird jedoch als gering eingeschätzt und muss im weiteren Verlauf näher beobachtet werden.

ZUSAMMENFASSUNG

Zusammenfassend lässt sich festhalten, dass bei einer hybriden Verwendung der Automatisierungsmöglichkeiten Workflow-Automation und Attended RPA ein hoher Automatisierungsgrad erreicht werden kann.

Im Falle des LT-Prozesses konnte zum einen der Gesamtaufwand für Mitarbeitende und somit die Kosten des Prozesses mit Hilfe der Automatisierung verringert werden. Zum anderen konnte in der Fallstudie eine Qualitätsverbesserung verzeichnet werden.

Jedoch ist auch zu vermerken, dass weiterhin der Einsatz eines Mitarbeitenden notwendig ist, da die Programme insbesondere bei der Interpretation von Daten und abweichenden Datenstrukturen an ihre Grenzen stoßen. Diese manuelle Überprüfung kann das Aktualisieren der Daten in den Systemen verzögern. Hierbei ist jedoch erneut anzumerken, dass im Rahmen dieses Projektes ausschließlich mit Attended RPA gearbeitet wurde und die Automatisierung somit auf ausschließlich strukturierte Daten beschränkt ist. Es wird daher für zukünftige Automatisierungen empfohlen, die Verwendung von KI-Technologien zu untersuchen, um das manuelle Eingreifen in den Prozess zu verhindern. Nichtsdestotrotz ist abschließend festzuhalten, dass bereits der Einsatz von Attended RPA das Automatisierungspotential geeigneter Office-Tätigkeiten deutlich erweitert und den Mitarbeitenden somit mehr Zeit für wertschöpfende und interessante Tätigkeiten schafft.

LITERATURVERZEICHNIS

Bardens Andrea (2020): Robotic Process Automation (RPA) in der DACH-Region. Analyse mit Blick auf Finance & Accounting. PwC. Online verfügbar unter <https://www.pwc.de/de/rechnungslegung/robotic-process-automation-rpa-in-der-dach-region.pdf>, Abruf am 01.05.2021.

Bygstad, Bendik (2017): Generative Innovation: A Comparison of Lightweight and Heavyweight IT. In: *Journal of Information Technology* 32 (2), S. 180-193.

Czarnecki Christian, Auth Gunnar (2018): Prozessdigitalisierung durch Robotic Process Automation. In: Barton Thomas, Müller Christian, Seel Christian (Hrsg.): *Digitalisierung in Unternehmen*. Angewandte Wirtschaftsinformatik. Springer Vieweg, Wiesbaden, Wiesbaden, S. 113-132.

Georgakopoulos, Diimitrios; Hornick, Mark; Sheth, Amit (1995): An overview of workflow management: From process modeling to workflow automation infrastructure. In:

Distrib Parallel Databases 3 (2), S. 119–153. DOI: 10.1007/BF01277643.

Gimpel Henner; Röglinger, Maximilian (2015): *Digital Transformation: Changes and Chances*. Project Group Business and Information Systems Engineering (BISE). Fraunhofer Institute for Applied Information Technology Fit. Augsburg/Bayreuth. Online verfügbar unter https://fim-rc.de/wp-content/uploads/2020/02/Fraunhofer-Studie_Digitale-Transformation.pdf Abruf am 10.06.2021.

Gimpel Henner; Röglinger, Maximilian (2015): *Digital Transformation: Changes and Chances*. Project Group Business and Information Systems Engineering (BISE). Fraunhofer Institute for Applied Information Technology Fit. Augsburg/Bayreuth. Online verfügbar unter https://fim-rc.de/wp-content/uploads/2020/02/Fraunhofer-Studie_Digitale-Transformation.pdf Abruf am 10.06.2021.

Koch, Christina; Fedtke, Stephen (2020): *Robotic Process Automation*. Springer Berlin Heidelberg, Berlin, Heidelberg.

Kroll, Christian; Dujak, Adam; Darius, Volker; Enders, Wolfgang; Esser, Marcus (2016): *Robotic Process Automation - Robots conquer business processes in back offices*. Capgemini Consulting. Online verfügbar unter <https://www.capgemini.com/consulting-de/wp-content/uploads/sites/32/2017/08/robotic-process-automation-study.pdf>. Abruf am 30.04.2021.

Langmann, Christian; Turi, Daniel (2020): *Robotic Process Automation (RPA) – Digitalisierung und Automatisierung von Prozessen*. Springer Fachmedien Wiesbaden, Wiesbaden.

Lacity, Mary; Willcocks, Leslie; Craig, Andrew (2015): The IT function and robotic process automation. In: *The Outsourcing Unit Working Research Paper Series*. (15/05), S. 1-39.

Lacity, Mary; Willcocks, Leslie (2016): A new approach to automating services. In: *MIT Sloan Management Review* 58 (1), S. 41-49.

Mohapatra, Sanjay (2013): *Business Process Reengineering: Automation Decision Points in Process Reengineering*; Springer Science + Business Media New York, New York.

Moritz, André; Rimbach, Felix (2006): *Soft Skills für Young Professionals: Alles, was Sie für Ihre Karriere brauchen*. GABAL Verlag GmbH, Offenbach.

Microsoft (2020): Types of process automation. Online verfügbar unter <https://docs.microsoft.com/en-us/power-automate/guidance/planning/various-types-process-automation>. Abruf am 19.06.2021

Osman, Cristina-Claudia (2019): Robotic Process Automation: Lessons Learned from Case Studies. In: *IE* 23 (4/2019), S. 66–71.

Oswald, Gerhard; Krcmar, Helmut (2018): *Digitale Transformation*. Springer Fachmedien Wiesbaden, Wiesbaden.

Penttinen, Esko; Kasslin, Henje; Asatiani, Aleksandre (2018): How to choose between Robotic Process Automation and Back-Ende System Automation. In: *Proceedings of the 26th European Conference on Information Systems Portsmouth (ECIS2018)*, UK.

Plattfaut, Ralf; Koch, Julian F.; Trampler, Michael; Conres, André (2020): PEPA: Entwicklung eines Scoring-Modells zur Priorisierung von Prozessen für eine Automatisierung. In: *HMD* 57 (6), S. 1111-1129.

Rashid, Naved; Ray, Saikat (2021): 3-steps-to-start-your-rpa-journey. Gartner. Online verfügbar unter <https://www.gartner.com/en/doc/733994-3-steps-to-start-your-rpa-journey>. Abruf am 30.04.2021.

Reinhardt, Kai (2020): *Digitale Transformation der Organisation*. Springer Fachmedien Wiesbaden, Wiesbaden.

Sikora, Sara; Hurley, Blythe; Tharakan, Anya George (2019): *Automation with intelligence. Reimagining the organisation in the 'Age of With'*. Deloitte. Online verfügbar unter

<https://www2.deloitte.com/content/dam/Deloitte/tw/Documents/strategy/tw-Automation-with-intelligence.pdf> Abruf am 30.04.2021.

Suhr, Frauke (2020): „Roboter Sind Auf Dem Vormarsch.“ Statista, Statista GmbH, Online verfügbar unter <https://de.statista.com/infografik/23800/anteil-der-industrieroboter-im-verarbeitenden-gewerbe/> Ab-ruf am 30.06.2021.

Syed, Rehan; Suriadi, Suriadi; Adams, Michael; Bandara, Wasana; Leemans, Sander J.J.; Ouyang, Chun et al. (2020): Robotic Process Automation: Contemporary themes and challenges. In: Computers in Industry 115, S. 103162.

Taulli, Tom (2020): The Robotic Process Automation Handbook. Apress Media LLC, Monrovia, CA.

Van der Aalst, Will M. P.; Bichler, Martin; Heinzl, Armin (2018): Robotic Process Automation. In: Business & Information Systems Engineering 60(4), S. 269–272

Willcocks Leslie; Hindle John; Lacity Mary (2019): Keys to RPA Success. How Blue Prism Clients are Gaining Superior Long-Term Business Value. Knowledge Capital Partners.

zukunftsInstitut (2021): Der Megatrend Konnektivität. zukunftsInstitut. Online verfügbar unter <https://www.zukunftsinstitut.de/dossier/megatrend-konnektivitaet/> Abruf am 24.06.2021

IT-SUPPORTED DATA ANALYSIS OF FUEL CONSUMPTION AND FLIGHT PROGRESS IN THE AVIATION SECTOR

Andreas Walter
Department of Industrial Engineering
Tor Vergata University of Rome
I-00133 Rome, Italy
E-Mail: walter.andreas@students.uniroma2.eu

KEYWORDS

Fuel Costs, Aviation Emission, Data Analytics, Excel, MATLAB.

ABSTRACT

Aviation continues to be an important means of transport for passengers and cargo. However, fuel savings and emissions reduction have become increasingly important in recent years. As part of a PhD thesis, possibilities for reducing fuel consumption by reducing the final reserve fuel were investigated. A smaller amount of fuel required leads to a reduction in the transported (fuel) weight and thus to a reduction in fuel consumption. Improved risk assessment, calculations based on better data (e. g. aircraft performance, route and weather) and better decision-making provide an enormous potential to optimise the amount of fuel needed without compromising safety levels. Proof of this must be provided with the help of suitable data collection, often with correspondingly large data sets. The statistical basis for this study was data over a five-year period provided by an airline from various reports, each with approx. 40 000 data sets. The assessment was carried out in several steps using various tools such as Excel or, in particular, MATLAB. A wide range of data on all aircraft movements is recorded in the internal database system of the air carrier which provided the information. By means of statistical analysis of existing fuel data, it was possible to prove that the planning and flight execution is reliable - as a representation of the basic level of safety. MATLAB provides a quick way to create clear and meaningful plots. With the help of this assessment, evidence of a performance-based approach to the approval of fuel planning procedure can be provided. This in turn can be used for the application of procedures requiring approval by the competent authority.

INTRODUCTION

Fuel costs are a driver for reducing the amount of fuel consumed. In recent years, however, increasing attention has also been paid to environmental impacts. Reducing the effect of global warming due to emissions has become an important target. An often-mentioned goal is to reduce global net aviation-created emissions

by 50% till 2050 compared to 2005 (IATA, 2019). As a result the reduction of emissions has become a major concern. Both aspects will be briefly discussed in the following.

Fuel Costs

Despite a downturn in 2020 and 2021, aviation is expected to recover to pre-pandemic levels. Already in 2015, around 3.5 billion passengers used air transport for their business and tourism purposes. This is an increase of around 6.4 % compared to 2014 (ICAO, 2016). For 2037, an IATA forecast predicts the number of air travellers reaching 8.2 billion (IATA, 2018). The industry is facing major challenges at the same time. Emissions from aviation, both domestically and internationally, account for about 2% of total global CO₂ emissions (ICAO, 2014). Fuel prices are linked to crude oil prices and are therefore volatile. Figure 1 (Airbus, 2014) shows the price development over a period of more than a decade. The fundamental upward trend can be seen, which has radically intensified in spring 2022.

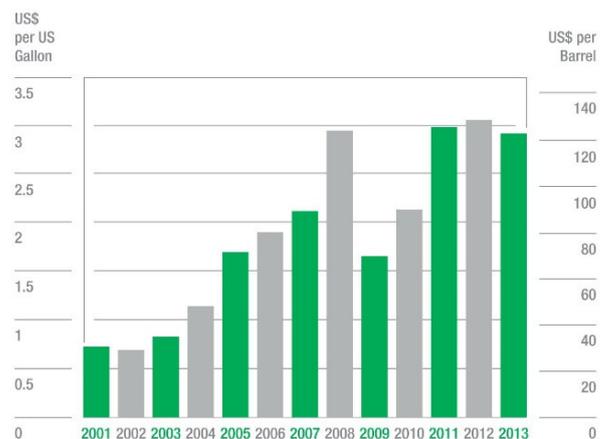


Figure 1: Monthly Jet Fuel Price Trend

The price is influenced by various factors such as political instability or high demand. Publications by International Air Transport Association (IATA) also show this trend, although a drop below \$35 per barrel can also be seen here in spring 2020. In spring 2022, prices have developed to over \$155 per barrel, according

the IATA Jet fuel price development (IATA, 2021). As a consequence fuel consumption and the associated costs represents a significant part of operating costs, see Table 1 (Airbus, 2014).

Table 1: Direct Operating Cost Breakdown

Cost factors	Share of costs [%]
Crew	18
Maintenance	25
Fuel	28
Navigation/Landing	24

Airline fuel expenditure varies from more than 30 per cent in 2012 to around 20 per cent in 2016, and has been projected to reach around 28.4 per cent of total expenditure in 2020. In 2021, it amounted to 19 per cent of total expenditure (Statista, 2022). As crude oil prices rise in 2022, the share of fuel costs is expected to move back towards the 2014 ratio.

Emission Reduction

The effects of global warming have been increasingly observed in recent years, leading to efforts to limit this effect. An often mentioned goal is to stabilise the temperature increase to below 2°C compared to pre-industrial levels (Intergovernmental Panel on Climate Change, 2015). 30% of all transport-related CO₂ emissions are caused by combustion (International Transport Forum, 2020). In 2022, almost every commercial aircraft is still equipped with a fossil fuel-based propulsion system. Alternative propulsion systems, such as the use of liquid hydrogen (LH₂), are intended to make aviation more environmentally friendly, but are not yet sufficiently available (Janić, 2014).

The most important emissions from aircraft are greenhouse gases and noise. The results of jet fuel combustion consist mainly of carbon dioxide CO₂, water vapor H₂O, methane and nitrous oxide NO_x. The Intergovernmental Panel on Climate Change (IPCC) notes that CO₂ and H₂O are simply the most common products of jet fuel combustion (Intergovernmental Panel on Climate Change, 1999). Their emission indices are 3.15 kg/kg fuel burned and 1.26 kg/kg fuel respectively (Intergovernmental Panel on Climate Change, 1999). For NO_x the IPCC constitutes that it is next most abundant engine emission. The range of NO_x emission is between 5 and 25 g NO_x per kg of fuel burned (Intergovernmental Panel on Climate Change, 1999). In addition to the emissions mentioned above, there are other factors such as water vapour, soot and sulphate aerosols and increased cloud cover due to the formation of contrails (Lee et al., 2020). Contrails, their formation and possible effects are an equal special part of aviation research. Fuel costs and the impact of

aviation on the environment have become increasingly important. In particular, reducing the impact of global warming due to emissions has become a goal. The commercial aviation industry has already developed and implemented many techniques to reduce fuel consumption for reasons of economy and cost effectiveness. However, further efforts are necessary.

FUEL SAVING VIA WEIGHT REDUCTION

An important consideration in optimising fuel consumption, whether for the future or current operations, is the weight of aircraft and equipment. It is the physics of flight, that for an aircraft to fly it must generate lift to overcome its weight. The generation of the required lift and the movement of the airframe through the air create drag. The engines generate the necessary thrust to overcome this drag and enable the movement to generate lift, see Figure 2 (Airbus, 2014).



Figure 2: Elementary forces on an airframe

Fuel consumption for a given route therefore depends, among other factors, on the weight of the aircraft. The more an aircraft weighs, the higher the fuel consumption. To keep fuel consumption as low as possible, it is most economical to carry only the minimum weight required for the route in question. Carrying more or even unnecessary weight increases the amount of fuel required and consumed in flight.

The fuel consumption, for carrying extra weight or extra fuel, is called Fuel Carriage Penalty (FCP). European Union Aviation Safety Agency (EASA) gives a value of about 3 % extra in fuel consumption per kg and flight hour for additional weight (EASA, 2016). International Civil Aviation Organization (ICAO) Doc 10013 gives a value of around 2.5 - 4.5 % additional fuel consumption, depending on the characteristics of the aircraft (ICAO 2014, p. 12). Ayra et. al mentions a value of up to 4 % additional fuel consumption per flight hour to cover additional weight. Thus, for a seven-hour flight that fills up with 5 000 kg of extra fuel, about 1 300 kg of this fuel is consumed just to carry it. For the aircraft investigated in the study, the factor is 0.0273 kg fuel per flight hour per kg additional weight.

As a result, excessive fuel consumption has a significant impact on a company's environmental impact and profit

(Ayra et al., 2014). Note: the economic influence of fuel prices, which play a role in the context of fuel tankering, is not considered here. Conversely, Ayra et al. mention the potential benefits of fuel savings. A 1% reduction in consumption would save about 100 tonnes of fuel per year for a medium-sized jet aircraft. This amount corresponds to an approximate annual cost reduction of €38 000 per aircraft (Ayra et al., 2014). The associated emission reduction for a medium-sized aircraft can be calculated from the above figures. At 100 tons of fuel, this would be 315 tonnes of CO₂, 126 tonnes H₂O and between 500 and 2 500 kg NO_x per aircraft per year. This effect clearly states the profound saving potential of a reduction of fuel unnecessarily being carried on board.

LEGAL FUEL REQUIREMENTS

Looking for aviation regulations, the ICAO is the first organisation that comes to mind on an international level. Founded in 1944, ICAO is a specialized agency of the United Nations. One of ICAO's objectives is to promote the safe and orderly development of international civil aviation. At the global level, ICAO establishes standards and recommended practices (SARPs) covering aviation safety, security, efficiency, economic development, and environmental protection. ICAO is developing the Global Aviation Safety Plan (GASP) and the Global Air Navigation Plan (GANP), both globally planned initiatives in the area of safety and air navigation services (ICAO, 2016). ICAO publishes various documents with different legal force. One are the Annexes to the Convention on International Civil Aviation in accordance with Article 37 of the Convention on International Civil Aviation. An Annex of significance to commercial aviation is Annex 6. Chapter 4 of Annex 6 contains requirements for flight operations, with subchapter 4.3 listing SARPs for flight preparation. Subchapters 4.3.6.1 and 4.3.6.2 require that, based on actual aircraft-specific data and operating conditions for the planned flight, a sufficient quantity of usable fuel is carried to safely accomplish the planned flight and to allow for deviations from planned operations (ICAO, 2018). ICAO SARPs are not directly binding, they have to be transposed into national law.

At the European Union (EU) level, the scenario is slightly different. Aviation regulations are developed by the EASA. EASA develops different levels of regulatory material. The EU itself knows different types of legal acts. Regulations, decisions and opinions are relevant for the aviation sector. Regulations have a binding effect throughout the EU. Decisions are binding for an EU country or an individual operator. EASA Opinions are not binding, they allow a statement (European Union, 2020). Regulation (EU) No. 965/2012 sets out the requirements and procedures for air operations.

Due to changes in Regulation (EU) No. 965/2012, which will apply from fall 2022, there was an adjustment in the fuel requirements. These were previously regulated in CAT.OP.MPA.150, among others, and will in future be referred to in CAT.OP.MPA.180 and the following sections (EASA, 2020b). Operators demonstrating certain capabilities will be able to use individual fuel schemes in the future. This is intended for operators who can demonstrate a defined safety level, thus reflecting the move towards performance-based regulations (EASA, 2020b). Data supporting the intended deviation is required to support the implementation of the individual fuel scheme. The Annex to Opinion No. 02/2020 already contains preliminary information on the draft Acceptable Means of Compliance (AMC) and Guidance Manual (GM) and information to be considered for the performance-based deviation. A non-exhaustive list of safety performance indicators (SPI) that can be used to measure safety performance are:

- flights with 100 % consumption of the contingency fuel;
- flights with a percentage consumption of the contingency fuel (e.g. 85 %), as agreed by the operator and the competent authority;
- difference between planned and actual trip fuel;
- landings with less than the final reserve fuel (FRF) remaining;
- flights landing with less than minutes of fuel remaining (e.g. 45 minutes), as agreed by the operator and the competent authority;
- 'MINIMUM FUEL' declarations;
- 'MAYDAY MAYDAY MAYDAY FUEL' declarations;
- in-flight replanning to the planned destination due to fuel shortage, including committing to land at the destination by cancelling the planned destination alternate;
- diversion to an en-route alternate (ERA) aerodrome to protect the FRF;
- diversion to the destination alternate aerodrome; and
- any other indicator with the potential to demonstrate the suitability or unsuitability of the alternate aerodrome and fuel planning policy (EASA, 2020a).

As can be seen from these indicators, airlines wishing to have an individual fuel scheme approved by the competent authority require a great amount of data and information on fuel consumption.

EVALUATION WITH MATLAB

In order to check the reliability of the fuel planning and execution of the flights under the current requirements, corresponding information must be evaluated over a sufficiently long period of time - with the corresponding

effort regarding the processing of the collected data. For further validation of the fuel data, real flight data of a globally operating European airline was chosen. For this purpose, data from a cargo airline was provided over a period of approximately 5 years (4 years and 10 months), from March 2016 to the end of December 2020. The aircraft used is a Boeing B777-200 in a cargo version. It has a maximum take-off mass of about 347 800 kg, a maximum landing mass of 260 800 kg and a dry operating weight of about 141 600 kg. This results in a maximum payload capacity of about 103 tonnes. The maximum fuel capacity for this version is about 144 024 kg. The route network served includes both major airports and some regional airports, resulting in a mix of short-, medium- and long-haul flights.

In the first step, the data was exported from the reporting system and analysed with Excel. This proved to be impractical for the reasons mentioned below. Therefore, MATLAB was used as a tool in the further steps.

Reporting System

For the data period provided, a large amount of flight information that can be evaluated via various reports was fed into the airline's data management system. This system collects and provides a variety of facts in different areas of interest, such as aircraft-related information, airport statistics, cosmic ray figures, crew data, on-time performance, and others. Flight information such as number of auto lands, information on city pairs, de-icing reports, delay reports, etc. are available as well. The data can be exported from the system in various formats, such as XML, Excel, PDF, MHTML, TIFF, Word or CSV.

With regard to fuel, the information is used, among other things, to compile statistics for the crews, for example as a decision-making aid for fuel planning. An electronic flight back (EFB) application is provided for this purpose. Specific fuel and flight information from 5 different reports was provided and analysable:

- Fuel Data Validation
- Dynamic Flight List
- Fuel Analyzer Reference List
- Fuel Analyzer App Data
- Flight List Uld Detail Load

In the first step of the analysis, the data was analysed with the help of Excel. For this purpose, the data was exported and corresponding statistical evaluations such as mean value, variance, etc. were carried out using Excel functions. This step was correspondingly time consuming. The examination of the 5 reports showed that they contain different quantities of data sets, respectively information. Missing data, wrong data or erroneous entries lead to such differences. Missing or

wrong data can for example result from connectivity issues of the used aircraft communications addressing and reporting system (ACARS), miscalculation due to absent input, accidental operation of the ACARS system or simply missing information. Lost information, which may result in the rejection of a dataset, can be: previous fuel, fuel uplift, density, off- or on block time, shutdown fuel.

The different number of data sets within the reports did not allow to merge the information in one report. Sometimes information where only partially or only in one report available, potentially leading to false information.

Data analyses consist of some standard components: Pre-processing, summarisation and visualisation modelling. The pre-processing of the data for MATLAB did not take place. Work was done directly with the exported data from the reporting system. Further investigations into, among other things, better data analysis with MATLAB are to be carried out, but are not part of this study. In the first step of the investigation, the visualisation possibilities of MATLAB were used.

In addition to the evaluation with Excel, which is time-consuming, the individual reports were therefore examined in MATLAB. For this purpose, the information from the reports, if applicable, was fed directly into MATLAB. For the analysis with MATLAB, the pure, unfiltered data from the reporting system was imported as a table via the Excel export used previously. No correction or adjustment of the data for incorrect, missing or erroneous entries took place in this step. Nevertheless, in the first step a quick evaluation was possible without the need for extensive MATLAB code. The following is an overview of the evaluations for the Dynamic Flight List Report and Fuel Analyzer App References List Report.

Dynamic Flight List

The Dynamic Flight List Report contains information on: date of flight, flight number, aircraft type, aircraft registration, departure airport, arrival airport, diversion information, scheduled and actual departure time (STD and ATD), scheduled and actual time of arrival (STA and ATA), expected departure time (ETD), expected arrival time (ETA), departure delay, airborne time, landing time, arrival delay, block time, flight time, scheduled time, flight status (scheduled, cancelled, departed, returned and arrived) and leg load. The Flight List Report contains information for 43 113 flights for the period 1st march 2016 till 31st December 2020.

The entries comprise 30 columns and 43 113 rows, i.e. a 43 113 x 30 matrix. Even after a brief examination of the data, almost 40 000 entries remained. 3 594 flights

where marked as cancelled, meaning they were planned but did not take place. They were only considered in the comparison of planned and flown sectors. The resulting 39 519 flights took place. Out of that, 52 flights (0.13 %) returned on block after going off block. None of these flights went airborne. 6 flights, out of the remaining 39 467 flights, could be identified as have to be diverted, which gives a diversion rate of 0.015 %. The following information was extracted out of the Flight List Report:

- 43 113 flights, average planned block time 7:02 hours, flight time 6:20 hours
- 3 594 flights cancelled, average planned block time 7:23 hours, flight time 0 hours
- 39 467 flights arrived, average planned block time 7:00 hours, actual block / flight time 6:49 / 6:20 hours
- 52 flights returned, average planned block time 7:32 hours, actual block time 26 minutes

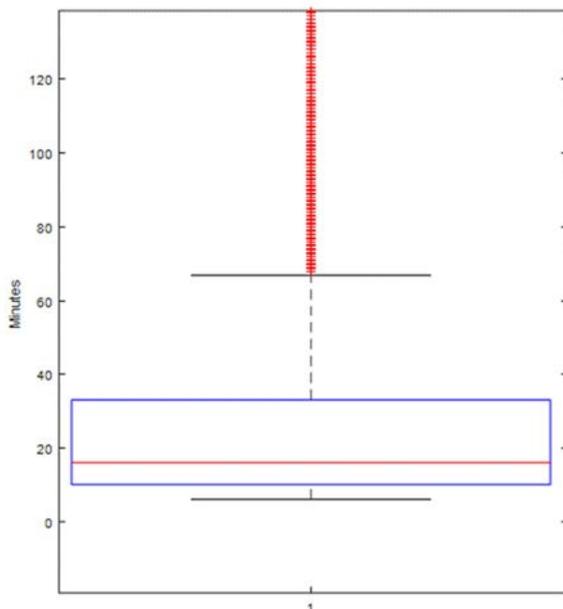


Figure 3: Arrival Delay

Delays in departures and arrivals are recorded in the Dynamic Flight List. Figure 3 shows a boxplot, created from the raw data over 43 113 flights, of arrival delays. Essential information on the boxplot are:

- Median: 16
- Maximum: 2 915
- Minimum: 6
- Number recorded: 43 113
- Finite Outliers: 703
- NaN or Inf: 36 915
- 75th percentile: 33

- 25th percentile: 10
- Upper adjacent: 67
- Lower adjacent: 6

As can be seen, the outliers drive the average upwards. Flights without delay have been recorded as Not a Number (NaN), in this case 36 915. In a more detailed analysis, special filters could be set to take a closer look at flights with delays of more than 60 minutes. The extreme value of 2 915 minutes should also be examined.

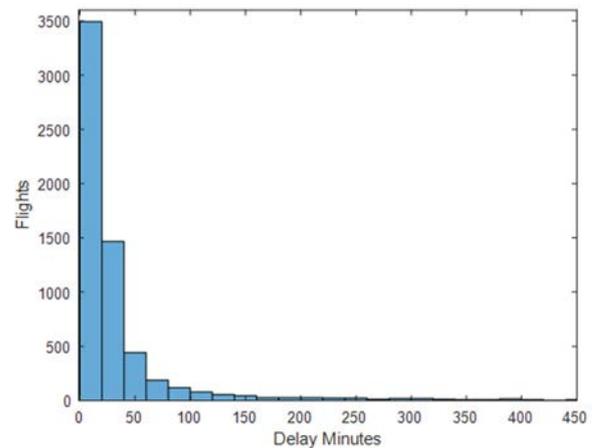


Figure 4: Delay Minutes

MATLAB offers the possibility to quickly create additional plots to the existing data. Delay information can also be displayed in a histogram for flights, where delays were recorded. Figure 4 shows the corresponding information. It can be seen that the majority of flights with delays had a delay of less than one hour. The sharp drop to the right shows the few flights with long delays. It should be mentioned that the delays do not refer exclusively to the actual course of the flight, but to the planned arrival. Delays in departure, e.g. due to weather or loading delays, lead directly to delays in arrival.

These two examples of illustrations show the advantages of a simple and fast extraction of information from the available data with the help of MATLAB.

In terms of fuel consumption, long flights are of interest, as a high proportion of fuel is required and transported as weight before consumption. This information can also be evaluated quickly. Figure 5 and 6 show the information on the distribution of block and flight times. Block time, as distinct from flight time, includes the time spent taxiing on the ground, before and after the flight. Evaluation with the help of boxplots of block and flight times show the relationship between these two parameters, however they are not listed here. The broad proportion of flights lasting more than five hours can be seen.

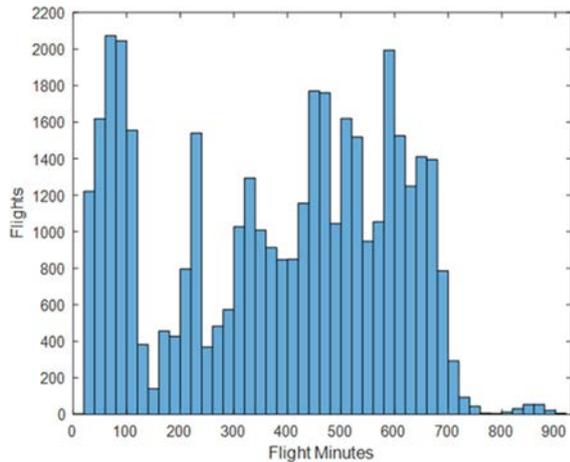


Figure 5: Flight Time

The evaluation of the flight times also shows a small but relevant number of flights with times between 800 and 900 minutes. These flights should be considered in more detail in the further development of the fuel consumption. Long flights, with high cargo and correspondingly high fuel content, offer the greatest opportunities for various parameters such as weather or flight route to have an influence.

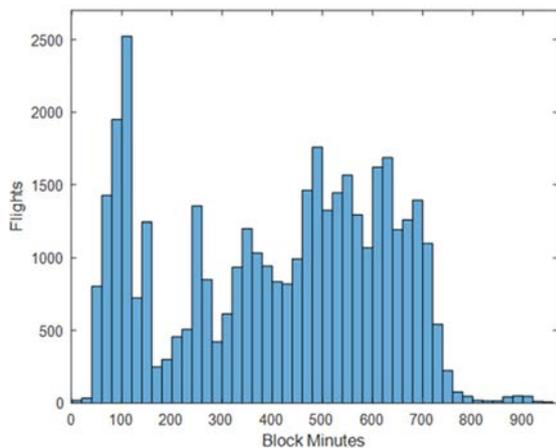


Figure 6: Block Time

The two peaks reflect the flights in the network. Short flights of 1 to 2 hours within Europe and medium to long haul flights worldwide.

Finally, information on the average load of the flights should be presented. As can be seen from Figure 7, the majority of flights were travelling with a load of between 60 and 100 tonnes, with an average of 70 tonnes. Some higher values than the maximum permissible loading, presumably incorrect values, were also found. This information may be relevant for the evaluation of routes without sufficient load. More interesting, however, are those flights that have a high load and possibly also a long flight time.

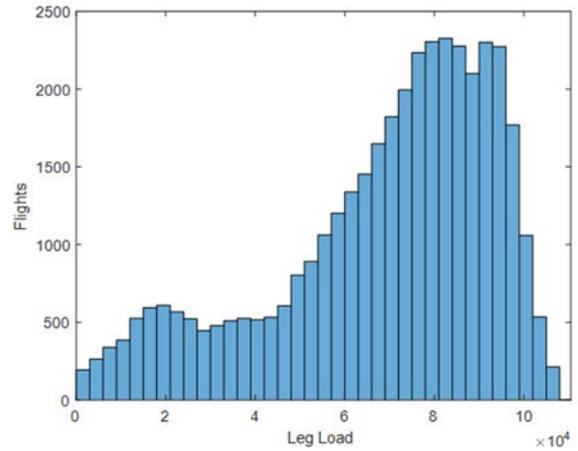


Figure 7: Leg Load

Further information, such as the city pairs flown and the aircraft's registration number, can be obtained from the above-mentioned report. However, these are not relevant for further fuel-related evaluation, in the first step.

Fuel Analyzer App References List Report

The Fuel Analyzer App References List Report contains information for 31 315 flights, period 1st march 2016 till end of December 2020. This reports is, amongst other things, fed by planning information and return information, send from the aircraft. It contains planned (p), actual (a) and corrected (c) information. Planned figures are self-explanatory, e. g. scheduled departure time, actual figures are true figures, like actual departure time. Actual figures describe the figures as sent by the aircraft. The report contains furthermore corrected figures. Corrected figures describe the difference between planned and feedback figures, based on the actual fuel decision and the payload of the flight. The Fuel Analyzer App References List comprises the data, which is used to feed the Fuel Analyzer App.

One of the safety performance indicators mentioned in the GM on Opinion 02/2020 is the number of landings with less than final reserve fuel (FRF). To obtain a statement about the planning and the progression of flights, a comparison of FRF and touch down fuel was done. Arriving at the destination, flights should have at least alternate fuel and final reserve fuel remaining on board or, in case of no alternate planning, FRF plus 15 minutes additional fuel. The final fuel reserve could be analysed in the Fuel Analyzer App References List report.

For 31 315 flights, the average FRF value was 2 954 kg, the mean, maximum, minimum and quantiles as shown in Figure 8. The maximum FRF value was recorded on the Shanghai - Frankfurt sector and the lowest on the Bahrain - Bangalore sector. Included here are 186

planned flights without alternate where an additional 15 minutes of fuel must be planned. These are included in the FRF of the reporting system of the air carrier, for technical reasons. Without these 186 flights, the average FRF would be 2 944 kg, only slightly below the overall average. These flights therefore have little influence and are only a small proportion of all flights. Figure 8 shows a corresponding boxplot evaluation.

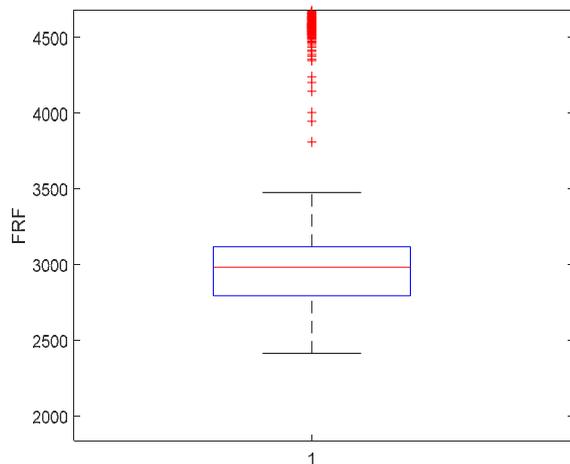


Figure 8: Final Reserve Fuel, with add. Fuel

The boxplot information on are:

- Median: 2 983
- Maximum: 5 031
- Minimum: 0
- Number recorded: 31 315
- Finite Outliers: 187
- NaN or Inf: 3 645
- 75th percentile: 3 118
- 25th percentile: 2 794
- Upper adjacent: 3 477
- Lower adjacent: 2 414

Interesting here is the flight with the entry "0 kg", which on closer inspection turned out to be a false entry. None of the analysed 31 315 flights landed below FRF fuel. The tightest buffer was around 300 kg of fuel at touch down above FRF for a flight which experienced extraordinary circumstances, more precisely the active eruption of a volcano, during the long-haul flight from China to America. Airspace was closed here. The affected flight was in close contact with the air traffic control centre throughout the entire flight and could thus make the decision to approach the destination airport.

The second plot, Figure 9, shows a histogram of the distribution of the FRF across the different flights . with a single pick at approx. 2 600 kg and the distribution around the mean of 2 954 kg.

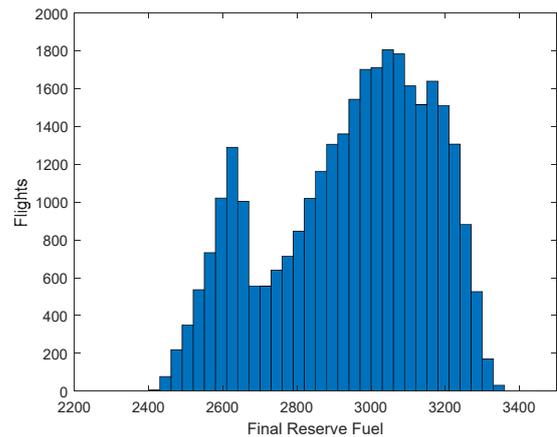


Figure 9: Final Reserve Fuel

CONCLUSION

The introduction of new regulations in the area of fuel requirements in the EU in 2022 enables air transport companies to optimise their planning and thus achieve savings in the area of fuel and emissions. At the same time, if a performance-based approach is chosen, proof of the equivalent level of safety must be provided before an approval can be granted. The evaluation of the necessary information, which is often available in large quantities, can be problematic. With the help of MATLAB, it is possible to quickly obtain clear representations of large and complex information out of the raw data. With the information obtained from the visualisations, proof of compliance with safety performance indicators can already be provided without any major pre-processing steps.

In the above section, only the first step of a statistical evaluation in this area was presented. However, this can already contribute to the demonstration of the safety performance indicators at this stage. Further evaluation steps can possibly provide even deeper and more comprehensive information which enable even better customised planning and also provide statistical information about specific routes or time periods. For this purpose, a more in-depth analysis of the data should be carried out with the help of MATLAB.

Further consideration of the interface to the reporting system is necessary, as there may be potential for further information and thus for optimisation. So far, only data from single point information, i. e. from planning or returned information, has been considered. No automatically transmitted data, which is transmitted at defined time intervals, was recorded. In addition, no evaluation of complex, time-stamped information took place, for which, for example, the use of the tall array would be necessary or possible.

AUTHOR BIOGRAPHIES



ANDREAS WALTER was born in Görlitz, Germany and started his career as a cadet in the German Air Force in 1997. As a young officer, he completed his studies in aerospace engineering at the University of the German Armed Forces in Munich. He then served as a technical officer in the German Air Force for seven years, dealing with all technical aspects of aviation. During this time, he studied business administration alongside his job at the FernUniversität in Hagen. After his military career, he moved to the German Civil Aviation Administration. Within the framework of a broad education, he got to know all the facets of aviation and eventually became an air operations inspector. He obtained the authorisation to fly various Airbus aircraft within the scope of commercial aviation, up to the authorisation as an instructor. His experience in supervisory work, his commercial and technical background provided the impetus and motivation to write a thesis in the field of fuel consumption. With the aim of combining his extensive and almost unique technical, organisational and operational experience, he is currently working on his PhD thesis at the Tor Vergata University of Rome and TH Wildau. His e-mail address is: walter.andreas@students.uniroma2.eu.

REFERENCES

- Airbus (2014) *Getting to grips with performance retention and fuel saving for A320 Family*. AIRBUS S.A.S., Blagnac Cedex.
- Ayra, E.S.; Insua, D.R. and Cano, J. (2014) To Fuel or Not to Fuel? Is that the Question? *Journal of the American Statistical Association* 109(506): 465–476, 10.1080/01621459.2013.879060.
- EASA (2016) *Notice of Proposed Amendment 2016-06 (A): Fuel planning and management*, Köln. See <https://www.easa.europa.eu/sites/default/files/dfu/NPA%202016-06%20%28A%29.pdf> (accessed 11/01/2020).
- EASA (2020a) *Appendix to Opinion No 02/2020: Draft Acceptable Means of Compliance (AMC) and Guidance Material (GM) to Annexes I-VIII to Commission Regulation (EU) No 965/2012*. See <https://www.easa.europa.eu/downloads/119252/en> (accessed 06/02/2022).
- EASA (2020b) *EASA Opinion No 02/2020: Fuel/energy planning and management*. See <https://www.easa.europa.eu/downloads/119248/en> (accessed 06/02/2022).
- European Union (2020) *Regulations, Directives and other acts*. See https://europa.eu/european-union/eu-law/legal-acts_en (accessed 25/04/2020).
- IATA (2018) *Press Release No: 62: IATA Forecast Predicts 8.2 billion Air Travelers in 2037*. See <https://www.iata.org/en/pressroom/pr/2018-10-24-02/> (accessed 19/04/2020).
- IATA (2019) *Aircraft Technology Roadmap to 2050*, Geneva. See <https://www.iata.org/contentassets/8d19e716636a47c184e7221c77563c93/technology20roadmap20to20205020no20foreword.pdf> (accessed 19/04/2020).
- IATA (2021) *Jet Fuel Price Monitor*. See <https://www.iata.org/en/publications/economics/fuel-monitor/> (accessed 13/12/2021).
- ICAO (2014) *Doc 10013: Operational Opportunities to Reduce Fuel Burn and Emissions*. International Civil Aviation Organization, Montréal.
- ICAO (2016) *Air Navigation Report*. International Civil Aviation Organization, Montréal.
- ICAO (2018) *Annex 6: Operation of Aircraft Part I – International Commercial Air Transport – Aeroplanes*, 11th edn. International Civil Aviation Organization, Montréal.
- Intergovernmental Panel on Climate Change (1999) *Aviation and the Global Atmosphere: Special Reports*. IPCC, Geneva.
- Intergovernmental Panel on Climate Change (2015) *Climate change 2014: Synthesis Report: Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Intergovernmental Panel on Climate Change, Geneva, Switzerland.
- International Transport Forum (2020) *The Carbon Footprint of Global Trade: Tackling Emissions from International Freight Transport*. See <https://www.itf-oecd.org/carbon-footprint-global-trade>.
- Janić, M. (2014) Greening commercial air transportation by using liquid hydrogen (LH₂) as a fuel. *International Journal of Hydrogen Energy* 39(29): 16426–16441, 10.1016/j.ijhydene.2014.08.011.
- Lee, D.S.; Fahey, D.W.; Skowron, A. et al. (2020) The contribution of global aviation to anthropogenic climate forcing for 2000 to 2018. *Atmospheric Environment*: 117834, 10.1016/j.atmosenv.2020.117834.
- Statista (2022) *Aviation industry - fuel cost 2011-2022*. See <https://www.statista.com/statistics/591285/aviation-industry-fuel-cost/> (accessed 02/04/2022).

Projekt Management Office

Ein Vorgehensmodell für die organisatorische Implementierung in KMUs

Leonie Holzheimer

Masterstudentin

Otto-Friedrich-Universität

Bamberg

Kapuzinerstraße 16

96047 Bamberg

E-Mail: leonie.holzheimer@web.de

Artur Rung

Leiter Produktmanagement &
PMO

xpecto AG

Ergoldinger Straße 2a

84030 Landshut

E-Mail: artur.rung@xpecto.com

Jürgen Wunderlich

Wirtschaftsinformatik

Hochschule Landshut

Am Lurzenhof 1

84036 Landshut

E-Mail: juergen.wunderlich@haw-landshut.de

ABSTRACT

Ein verstärktes Projektaufkommen stellt heutige Unternehmen häufig vor die Herausforderung, Projekte innerhalb eines vorgegebenen Zeitrahmens erfolgreich abzuschließen. Um diesen Herausforderungen strukturiert entgegenzuwirken, ist eine Priorisierung von Projekten notwendig, wobei Ansätze eines Multiprojektmanagements verwendet werden können. Für die Umsetzung des Multiprojektmanagements sowie der Standardisierung des Projektmanagements wird zumeist eine separate Abteilung in Unternehmen geschaffen, sodass eine unternehmensweite Projektplanung- und Steuerung umgesetzt werden kann. Diese Abteilung wird oft Projekt Management Office genannt und umfasst ferner die Aufgabe der strategischen Ausrichtung der aufkommenden Projekte. Die Einführung eines Projekt Management Offices ist firmenspezifisch zu gestalten, kann jedoch mittels vordefinierter Konzepte unterstützt werden. Dahingehend wird in diesem Artikel ein Vorgehensmodell zur organisatorischen Implementierung eines Projekt Management Offices in KMUs entwickelt und dessen Anwendung am Beispiel eines Softwareentwicklungsunternehmens beschrieben, das ein professionelles Einzelprojektmanagement sowie ein standardisiertes Projektportfoliomanagement etabliert.

SCHLÜSSELWÖRTER

Projektmanagement, Einzelprojektmanagement, Projektportfoliomanagement, Projektportfoliocontrolling, Gesamtprojektplanung, PMO, Change Management, KMU

BEDEUTUNG EINES PMO

Wie eine Studie zeigt, besteht schon seit mehreren Jahren insbesondere im informationstechnischen Bereich ein zunehmendes Interesse an der Einführung eines Projekt Management Offices (Frost et al. 2014). Dies ist auf die Komplexität von IT-Projekten zurückzuführen, bei denen gegenseitige Projektabhängigkeiten und unvorhersehbare Wechselbeziehungen die Ursache für eine erschwerte Projektplanung darstellen (Pawellek und Schramm 2017). Eine abteilungsübergreifende Projektplanung, um Interdependenzen zwischen unterschiedlichen Projektvorhaben berücksichtigen zu können, kann ohne zentrale Organisationseinheit nicht gewährleistet werden (Wyrozebski 2015). Der Verantwortungsbereich der einzelnen Projektleiter liegt meist nur auf deren zugewiesenen Projekten, sodass sie keinen Blick auf das gesamte Projektportfolio werfen (Lomnitz 2008). Ein professionelles Projektmanagement, das mittels eines Projekt Management Office (PMO) umgesetzt werden kann, stellt eine adäquate Methode dar, um Projekte zukünftig effizient organisieren und ohne Terminverzögerungen abschließen zu können.

Der Aufbau eines PMO lässt sich vielfältig gestalten, wobei die Unternehmensgröße und Organisationsstruktur

ausschlaggebend sind (Frost et al. 2014). Es empfiehlt sich, um eine möglichst effiziente Organisationseinheit zu schaffen, den Aufbau eines firmenindividuellen PMO anzustreben. Wie eine Studie aus dem Jahr 2020 ergeben hat, erfordern besonders klein- und mittelständische Unternehmen eine Projektmanagement-Methode, die auf den Mensch ausgerichtet ist und Flexibilität bietet (Parizotto et al. 2020). Dahingehend fokussiert sich dieser Beitrag auf die Bereitstellung eines Vorgehensmodells zur Implementierung eines auf die unternehmensspezifischen Anforderungen ausgerichteten PMOs in klein- und mittelständischen Unternehmen.

DER WEG ZUR IMPLEMENTIERUNG EINES PMOs

Das Vorgehensmodell enthält die vier grundlegenden Phasen *Ist-Analyse*, *Konzeptionierung*, *Implementierung* und *Übergang in den Regelbetrieb*, die während des gesamten Transformationsprozesses schrittweise zu durchlaufen sind (vgl. Abb. 1). Es sei jedoch angemerkt, dass parallel zu diesem vierstufigen Verfahren auch die Erfolgsfaktoren des Stakeholder Managements und des Change Managements zu berücksichtigen sind. Ein erfolgreiches Stakeholder Management zeichnet sich durch eine frühzeitige Miteinbeziehung von Stakeholdern aus, um so potenzielle Unterstützer und/oder Blockierer zu identifizieren und letztendlich einen reibungslosen und erfolgreichen Veränderungsprozess gewährleisten zu

können. Ebenfalls spielt während des gesamten Prozesses zur Einführung eines unternehmensweiten professionellen Projektmanagements das Change Management eine wichtige Rolle. Wie oftmals beobachtet wurde, sind



Abbildung 1 - Vorgehensweise zur Etablierung eines PMOs

Veränderungen im gewohnten Umfeld nicht gerne gesehen, sodass bei Optimierungen oder einer Neugestaltung von gewohnten Abläufen in der Regel schnell auf Widerstand seitens der Stakeholder gestoßen wird (Pfetzing und Rohde 2017). Ebenfalls hat, wie die Forschung von John P. Kotter ergeben hat, der Rückfall in gewohnte Verhaltensmuster eine beeinträchtigende Wirkung auf den Erfolg bzw. Misserfolg von Projekten (2011). Aus diesem Grund definierte Kotter das 8-Phasen-Modell, welches Projektbeteiligte während des gesamten Veränderungsprozesses begleiten soll, um ein positives und nachhaltiges Resultat erzielen zu können.

Ist-Analyse

Die erste Phase zur Entwicklung und letztlich Einführung eines PMOs-Konzeptes beinhaltet die Ist-Analyse des Unternehmens. Bei der Bestimmung des Status quo sind eine Vielzahl von Aspekten zu berücksichtigen, welche sich in die folgenden drei Kernelemente der Ist-Analyse zusammenfassen lassen: (1) Projektmanagement Prozesse & Methoden, (2) Projektorganisation & Unternehmensrichtlinie und (3) Projektmanagement Kompetenzen & Leitvorstellungen (vgl. Abb. 2).

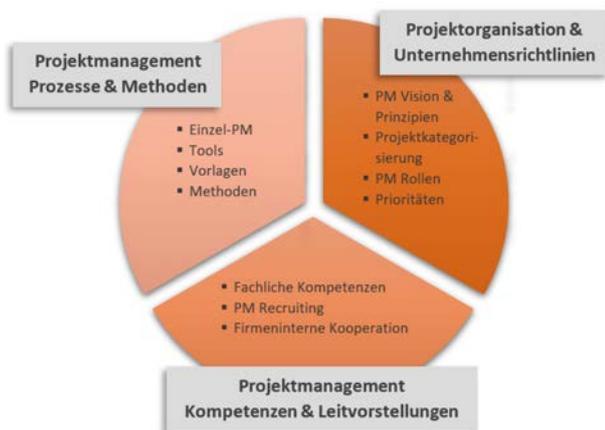


Abbildung 2 - Elemente der Ist-Analyse

Konzeptionierung

In Folge der evaluierten Ausgangssituation des Unternehmens erfolgt die Konzeptionierung eines firmenspezifischen PMOs. Darin sollen die Ziele sowie die Größe

des zukünftigen PMOs bestimmt werden. Die nachhaltige Etablierung eines PMOs lässt sich am besten durch eine Anpassung der Aufbau- und Ablauforganisation umsetzen (Sandrino-Arndt und Amberg 2010). Das Konzept sieht daher vor, das PMO in die Unternehmensstruktur zu verankern, indem es in der gewünschten/notwendigen Hierarchieebene eingegliedert wird. Einer Umfrage zu Folge erfolgt die Verankerung eines PMOs überwiegend in Form einer Stabstelle oder durch die Einbindung in die Linienorganisation. Dabei wurde außerdem festgehalten, dass bei größeren Unternehmen die Wahl überwiegend auf die Verankerung des PMOs in die Linienorganisation fällt (Frost 2014). Demgegenüber ist aufgrund knapper Ressourcen gut vorstellbar, dass das PMO bei KMUs im Idealfall als Stabstelle einzugliedern ist, die nur mit einem Mitarbeiter besetzt wird. Während in der Aufbauorganisation die organisatorische Eingliederung des PMOs bestimmt wird, wird mit der Ablauforganisation die Definition des zukünftigen Aufgabenbereichs des PMOs in Verbindung gebracht. Eine Pyramide, die sich aus den aufeinander aufbauenden Ebenen Einzelprojektmanagement, Operatives Projektportfoliomanagement und Strategisches Projektmanagement zusammensetzt (vgl. Abb. 3) ist als ein kompaktes Gesamtpaket eines PMOs zu betrachten. In Abhängigkeit vom Kompetenz-

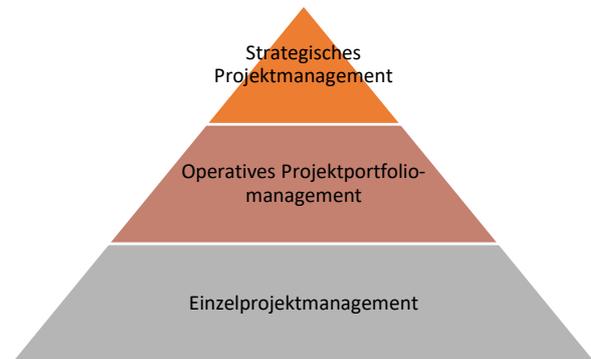


Abbildung 3 - Elemente des Projekt Management Offices

profil der Projektleitung, dem Auslastungsgrad der Mitarbeiter sowie dem Umfang des zukünftigen Aufgabenportfolios des PMOs, ist zu bestimmen, in welcher Ausprägung die einzelnen Elemente etabliert werden sollen (Ortner und Stur 2019). Der Grundstein der Pyramide wird durch die Definition eines einheitlichen Einzelprojektmanagement gelegt, das standardisierte Projekt-abläufe vorsieht. In der zweiten Ebene der Pyramide ist das Operative Projektportfoliomanagement angesiedelt, das sich primär auf das Projektportfoliocontrolling konzentriert. Darüber hinaus legt das Strategische Projektmanagement bzw. Strategische Projektportfoliomanagement das Hauptaugenmerk auf die strategische Projekt- und Ressourcenplanung.

Nun stellt sich die Frage, wie am besten vorzugehen ist, um ein unternehmensindividuelles Vorgehensmodell für das Einzelprojektmanagement zu entwerfen. Zunächst sollte ein einheitliches Verständnis für die Thematik des Projektmanagements geschaffen werden, um mögliche Konflikte vorzubeugen. Jedoch stellt ein standardisiertes

Projektmanagement sowohl Chancen als auch Risiken (oder Herausforderungen) für Projektbeteiligte dar. Dahingehend gilt es herauszufinden, wie starr ein Projekt- ablauf gestaltet werden kann, ohne die Projektarbeit un- nötig zu erschweren. Zu den bekanntesten Formen des Einzelprojektmanagements zählen das traditionelle, agile oder hybride Vorgehensmodell. Um nun ein adäquates Vorgehensmodell identifizieren zu können, das auf die Unternehmenssituation abgestimmt ist, sind einige As- pekte zu beachten, die zur Entscheidungsfindung heran- gezogen werden können:

- Größe und Kompetenzen des Teams
- Standortübergreifende/räumliche Strukturen des Teams
- Beständigkeit der Anforderungen
- Komplexitätsgrad der Projekte
- Zusammenspiel von Linien- und Projektaktivitäten
- Anforderungen und Vorgaben vom Kunden (Timinger 2017)

Selbst nach der Auswahl bzw. Einführung einer unter- nehmensweiten allgemeinen Vorgehensweise im Pro- jektmanagement gilt es zu beachten, dass das erarbeitete Konzept stets evaluiert und ggf. überarbeitet werden sollte. Dabei kann nach dem Kontinuierlichen Verbesse- rungsprozess (KVP) nach Deming mit den vier Phasen Plan – Do – Check – Act vorgegangen werden (Chakraborty 2016).

Nachdem ein einheitliches Konzept für das Einzelpro- jektmanagement entwickelt wurde, wird zur Realisierung des zweiten und dritten Bausteins der PMO-Pyramide die Etablierung eines Portfoliomanagement angestrebt. Da- für wird eine fünfstufige Handlungsanleitung zur kon- zeptionellen Entwicklung eines Portfoliomanagements bereitgestellt (vgl. Abb. 4).

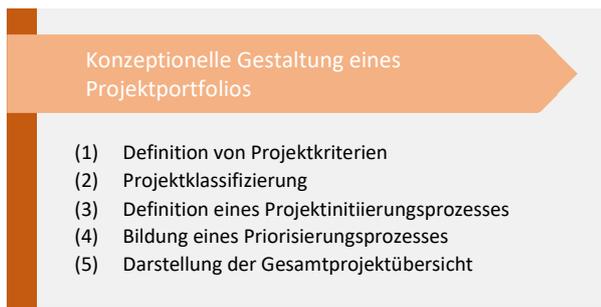


Abbildung 4 - Entwicklung eines Portfoliomanagements

Die im ersten Schritt vorzunehmende Definition von Pro- jektkriterien gewährleistet die Abgrenzung der Projekttä- tigkeiten von den Linientätigkeiten, um nachfolgend die einzelnen Projekte klassifizieren zu können. Im dritten Schritt erfolgt die Festlegung eines Projektinitiierungs- prozesses, der eine vollständige Erfassung aller Projek- tanträge im Projektportfolio sicherstellen soll, um in der Folge ein effektives Portfoliomanagement ermöglichen zu können.

Zur Gestaltung eines solchen Prozesses können z.B. fol- gende Fragestellungen herangezogen werden:

- Wie erfolgt die Sammlung von Projektideen und -anträgen?
- In welchem zeitlichen Rhythmus erfolgt eine Aktualisierung des Projektportfolios?
- Wie und durch wen erfolgt eine Genehmigung von Projektanträgen?
- Befinden sich alle Beteiligten auf einer gemeinsa- men Informationsbasis?

Der vierte Schritt des Konzeptes sieht die Definition ei- nes Priorisierungsprozesses vor, indem die Phasen De- zentrale Planung – Zentrale Vorhabensplanung und Res- sourcenzuteilung zu durchlaufen sind. Der letzte Schritt, die Darstellung der Gesamtprojektübersicht, zielt auf die Bereitstellung eines zentralen Informationsspeichers, der die abteilungsübergreifenden Projektplanung unterstüt- zen soll.

Implementierung

Nachdem das gesamte Konzept des PMOs erarbeitet wurde, ist nun die Implementierung zu vollziehen. Diese sollte möglichst nur unter Vorbereitung der Mitarbeiter sowie in schrittweisen Etappen realisiert werden. Die Im- plementierungsphase beinhaltet außerdem die Verdeutli- chung des Nutzens eines PMO für die betroffenen Stake- holder. Bei Berücksichtigung der zuvor genannten Punkte, steht der anschließenden Implementierung des PMOs nichts mehr im Wege.

Übergang in den Regelbetrieb

Der Übergang in den Regelbetrieb leitet sozusagen den vorübergehenden Abschluss des Veränderungsprojektes bzw. Transformationsprozesses ein. Hierbei sollte der Aufgabenbereich des PMOs genauestens eingegrenzt wer- den. Weiterhin ist der PMO-Mitarbeiter entsprechend zu schulen. Außerdem ist nochmals darauf hinzuweisen, dass trotz erfolgreicher Etablierung einer zentralen Orga- nisationseinheit eine kontinuierliche Weiterentwicklung des gesamten Konzeptes angestrebt werden sollte.

DIE IMPLEMENTIERUNG EINES PMOs IM UNTERNEHMEN XPECTO AG

In Folge des Unternehmenswachstums musste über orga- nisatorische Anpassungen nachgedacht werden. Die zu- nehmende Anzahl von Projektaufträgen hat bei der xpecto AG dazu geführt, dass Projekte mittels der ge- wohnten Arbeitsweise (ohne standardisiertes Projektma- nagement) nicht mehr effizient bewältigt werden konn- ten.

Besonders bei Organisationen mit einem hohen Projekt- aufkommen stellt ein professionelles und unternehmens- weites Projektmanagement ein wirkungsvolles Mittel für eine effiziente Projektabwicklung dar. Auch im projekt- orientierten Softwareentwicklungsunternehmen xpecto AG wurde diesbezüglich Handlungsbedarf erkannt. Auf- grund des Unternehmenswachstums entstand der

Wunsch nach einem professionellen Projektmanagement, das mittels eines PMOs umgesetzt werden sollte. In Anlehnung an das vierstufige Vorgehensmodell zur Implementierung eines PMOs wurde dieser Wunsch verwirklicht.

Zunächst wurde durch die *Ist-Analyse* der Projektorganisation das Problem eines nicht vorhandenen Portfoliomanagements identifiziert, das erfolgskritisch für eine effiziente Projektplanung ist. Der Wille einer Projektpriorisierung und die dementsprechende Umsetzung dieser Projekte konnten somit nicht erfüllt werden. Darüber hinaus wurde auch das Nichtvorhandensein und der Bedarf nach einem unternehmensweiten standardisierten Projektmanagement-Vorgehen festgestellt. Dieses soll den vorherig unstrukturierten Projektablauf entgegenwirken und überflüssigen Arbeitsaufwand durch Bereitstellung einheitlicher Verfahren auf ein Minimum senken. Die Projektmanagement-Kompetenzen der einzelnen Projektleiter wurden durch eine zertifizierte Projektmanagement-Schulung verbessert, sodass ein einheitliches Verständnis für dieses Fachgebiet vorhanden ist und somit darauf aufbauen ein Konzept für ein erfolgreiches PMO entworfen werden konnte.

In der *Konzeptionierungsphase* wurde zu Beginn die organisatorische Eingliederung des PMOs in Angriff genommen, wobei der Wunsch nach der Integration als Stabsstelle verfolgt wurde. Das PMO soll in dieser Position den strategischen Ambitionen der Geschäftsleitung folgen und gleichzeitig als unterstützendes Instrument zur abteilungsübergreifenden Projektplanung für Projektbeteiligte dienen. Das geplante Aufgabenspektrum des PMOs wird durch nachfolgende Abbildung (vgl. Abb. 5) veranschaulicht.



Abbildung 5 - Aufgaben PMO

Wie Abbildung 5 zeigt, soll das zukünftige PMO der xpecto AG alle fünf Aufgabenfelder eines PMOs übernehmen. Darunter fällt das Strategische Projektmanagement, das Projektportfoliocontrolling, die Entwicklung und Weiterentwicklung der Projektmanagement-Methodik (z.B. durch die Definition von Standards), das Operative Projektcontrolling sowie Fortbildung durch Training und Coaching.

Zur Durchführung dieser Aufgaben wurden Kern- und Unterstützungsprozesse definiert, um die exakten Tätigkeiten des PMOs festzulegen. Die Unterstützungsprozesse beziehen sich dabei auf das Operative Projektcontrolling, das eine Unterstützung für die Projektplanung und -umsetzung bezwecken soll, z.B. durch Bereitstellung eines Leitfadens zur Projektabwicklung sowie standardisierter Arbeitsmittel. Ein weiterer Punkt, der im Rahmen der Unterstützungsprozesse definiert wurde, befasst sich mit dem Thema Fortbildung. Hierbei wird die Planung und Koordination von internen und externen Weiterbildungsmöglichkeiten betrachtet.

Bei der Entwicklung eines Vorgehensmodells für das Einzelprojektmanagement wurde ein unternehmensindividuelles Konzept erarbeitet, das zunächst für Großprojekte entwickelt wurde und anschließend entsprechend an die restlichen Projektarten angepasst wurde. Es verfolgt die klassische Herangehensweise mit den fünf Phasen Initialisierung, Definition, Planung, Steuerung und Abschluss, wobei die Steuerungsphase als agile Phase ausgearbeitet wurde, um eine gewisse Flexibilität mittels der Scrum-Methodik im Projektverlauf ermöglichen zu können. Das fünfstufige Vorgehensmodell diente auch bei der konzeptionellen Gestaltung des Projektportfolios als Orientierung. Hier wurden zunächst Projektkriterien definiert, um anschließend anhand dieser Kriterien die Ausprägung der Projektaktivitäten bestimmen zu können. Als Kriterien wurden bspw. die Laufzeit und die inhaltliche Komplexität eines Vorhabens sowie die Größe des gesamten Projektteams berücksichtigt. Daraus resultierend erfolgte eine Klassifikation der Projekte in Mini-Projekt, mittleres Projekt oder Groß-Projekt (vgl. Abb. 6).

Kriterien	Ausprägung	
Größe des gesamten Projektteams	klein	groß
Dauer	dauerhaft	sehr lang
Inhaltliche Komplexität	niedrig	hoch
Neuartigkeit für das Projektteam	gering	hoch
Qualitätsrisiko	keines	sehr hoch

Kein Projekt	Mini-Projekt	Mittleres-Projekt	Groß-Projekt
--------------	--------------	-------------------	--------------

Abbildung 6 - Projektklassifikation

Für den Projektinitiierungsprozess wurde ein Ablauf definiert, der Projektanträge bis hin zu Umsetzungsentscheidungen abbilden soll (vgl. Abb. 7).

Hierbei liegt der Fokus auf der vollständigen Erfassung von Projektideen und -anträgen mittels eines zentralen Systems, wodurch eine effiziente und abteilungsübergreifende Zusammenarbeit ermöglicht werden soll. Die Bildung eines Priorisierungsprozesses zielt auf eine zukünftige Projektplanung nach Prioritäten, zur optimalen Steigerung des Unternehmenswertes, ab. Der Prozess beinhaltet die drei Phasen Dezentrale Planung, Zentrale Vorhabensplanung und die Ressourcenallokation.

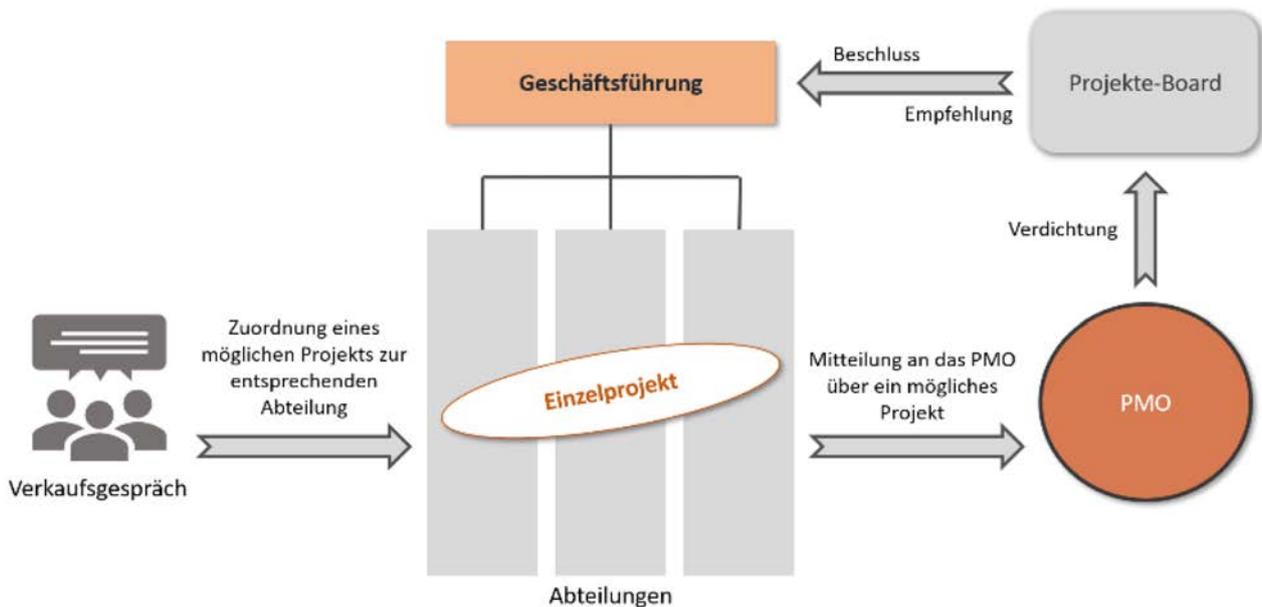


Abbildung 7 - Projektinitiationsprozess

Die Idee ist, dass zunächst in der dezentralen Phase Projektleiter ihre Projekte innerhalb der jeweiligen Produktgruppe nach selbst gewählten Priorisierungskriterien einstufen, wobei als Priorisierungskriterium z.B. die Qualitätsverbesserung eines Softwareprodukts im Vordergrund stehen kann.

Im zweiten Schritt ist das Ziel die einzelnen dezentralen Priorisierungspläne der unterschiedlichen Produktgruppen zu bündeln, um einen zentralen Vorhabensplan aufstellen zu können. Zur Vereinigung der einzelnen Pläne wird das Verfahren der Blockpriorisierung angedacht.

Das genaue Vorgehen zur Priorisierung mittels Blockbildung kann im Buch „Projektportfolio-Management“ von Hirzel et al. (S. 97f) nachgelesen werden.

Nach Durchführung der Blockpriorisierung sind nachträglich noch Interdependenzen zwischen Projekten zu prüfen und ggf. zu berücksichtigen. Nach Abschluss einer zentralen Ressourcenplanung ist zuletzt die Phase der Ressourcenallokation zu beachten. Hierbei werden die zur Verfügung stehenden Ressourcen den entsprechend priorisierten Projekten zugeteilt bis das gesamte Kontingent aufgebraucht ist. Zur Darstellung der Gesamtprojektübersicht setzt die xpecto AG auf eine firmeneigene Software-Lösung, die zukünftig unternehmensweit als Standard-Projektmanagementtool genutzt werden soll. Ziel ist es, eine gemeinsame Datenbasis durch ein zentrales System zu schaffen und Projekte unkompliziert und übersichtlich aufeinander abstimmen zu können.

Nach der Konzeptionierung des PMOs wurde der Schwerpunkt der *Implementierungsphase* auf das Einzelprojektmanagement gelegt und das Projektportfoliomanagement als zweitrangig betrachtet, um der Idee einer schrittweisen Implementierung gerecht zu werden.

Trotz *Übergang in den Regelbetrieb* und dem damit verbundenen Abschluss des Veränderungsprojekts, ist eine kontinuierliche Weiterentwicklung der erarbeiteten Methoden und Standards der xpecto AG zu empfehlen. Mögliche Schwachstellen werden erst durch die Anwendung von Konzepten ersichtlich, sodass laufende Optimierungen im Sinne eines kontinuierlichen Verbesserungsprozesses sinnvoll sind.

FAZIT

Zusammenfassend lässt sich festhalten, dass Probleme im Projektmanagement-Umfeld mit Hilfe eines PMOs vermieden werden können. Denn ein PMO, das für die Planung einer bereichsübergreifenden Projektarbeit zuständig ist, stellt eine enorme Erleichterung für das gesamte Projektmanagement dar. Voraussetzung dafür ist ein auf die Organisationsstruktur abgestimmtes und mit den Stakeholdern vereinbartes Konzept, um ein bestmögliches Ergebnis für das Unternehmen erzielen zu können. Als Leitfaden zur Implementierung eines unternehmensspezifischen PMOs könnte dafür das im Artikel vorgestellte Vorgehensmodell dienen. Dieses wurde bei der xpecto AG erfolgreich im Rahmen eines internen Organisationsprojekts angewandt. Der Aufbau des Modells mit den Phasen Ist-Analyse, Konzeptionierung, Implementierung und Übergang in den Regelbetrieb hat sich dabei als sinnvoll erwiesen. Bereits im ersten Jahr seiner Einrichtung als Stabsstelle hat sich das PMO als fester Bestandteil innerhalb der Organisation etabliert und sorgt über viele Geschäftsprozesse hinweg für erfolgreiche Projekte.

LITERATUR

- Chakraborty, A. 2016. Importance of PDCA Cycle for SMEs. In: SSRG International Journal of Mechanical Engineering (SSRG - IJME) – Volume 3 Issue 5 – May 2016
- Frost, R.; Klug, T.; Henning, A.; Ehry, R.; Götz, K. 2014. Das PMO in der Praxis. Verbreitung - Akzeptanz - Erfolgsmessung. In: Empirische Studie 2013/2014 - GPM Deutsche Gesellschaft für Projektmanagement e.V. KNOW-HOW
- Hirzel, M.; Alter, W.; Niklas, C. 2019. Projektportfolio-Management, Wiesbaden
- Kotter, J. P. 2011. Leading Change. Wie Sie Ihr Unternehmen in acht Schritten erfolgreich verändern, München
- Lomnitz, G. 2008. Multiprojektmanagement: Projekte erfolgreich planen, vernetzen und steuern, München
- Ortner, G. und Stur, B. 2019. Das Projektmanagement-Office, Berlin/Heidelberg
- Parizotto, L.; Tonso, A.; Monteiro de Carvalho, M. 2020. The challenges of project management in small and medium-sized enterprises: a literature review based on bibliometric software and content analysis. *Gestão & Produção*, 27(1)
- Pawellek, G. und Schramm, A. 2017. Komplexe IT-Projekte effizient managen. Klassisches Projektmanagement um innovative Methoden ergänzen. In: it-daily.net
- Pfetzinger, K. und Rohde, A. 2017. Ganzheitliches Projektmanagement, Gießen
- Sandrino-Arndt, B. und Amberg, M. 2010. Handbuch Project Management Office. Mit PMO zum strategischen Management der Projektlandschaft, Düsseldorf
- Timinger, Holger. 2017. Modernes Projektmanagement. Mit traditionellem, agilem und hybridem Vorgehen zum Erfolg, Weinheim
- Wyrozowski, P. 2015. PMO in Focus of Management Sciences. In: Kommunikation in multikulturellen Projektteams, Frankfurt am Main

KONTAKT



LEONIE HOLZHEIMER ist in München geboren und absolvierte 2020 an der Hochschule Landshut ihr Bachelorstudium in Wirtschaftsinformatik. Ihre Abschlussarbeit zum Thema PMO schrieb sie in Zusammenarbeit mit der xpecto AG. Nach ihrem Bachelorabschluss begann sie ihr Masterstudium an der Universität Bamberg, das sie voraussichtlich im Herbst 2022 abschließen wird. Ihre E-mail Adresse ist: leonie.holzheimer@web.de



ARTUR RUNG verantwortet den Bereich Produktmanagement & PMO bei der xpecto AG. Seine berufliche Leidenschaft liegt dabei an der Entwicklung von Produkten, der strategischen Optimierung von Prozessen und dem Projektmanagement. Seine E-Mail Adresse ist: artur.rung@xpecto.com



JÜRGEN WUNDERLICH ist seit 1. Oktober 2010 Professor und Studiengangsleiter für Wirtschaftsinformatik an der Hochschule Landshut. Dort entwickelt er v.a. Methoden zur wertschöpfungsorientierten Optimierung betrieblicher Strukturen und Prozesse in Produktion und Logistik. Darüber hinaus führt er als interdisziplinärer Prozessoptimierer laufend Projekte in der Fertigungs- und Logistikindustrie durch und verfügt außerdem über eine langjährige Operations-Erfahrung bei einer international führenden Top-Management-Beratung. Seine E-Mail Adresse ist: juergen.wunderlich@haw-landshut.de

Modifizierung und Verbesserung des Softwarelizenzmanagements durch den Einsatz von Managed Services

Giovanni Ernst

Technische Hochschule
Mittelhessen

Fachbereich MND
Wilhelm-Leuschner-Str. 13
61169 Friedberg
giovanni.ernst@mnd.thm.de

Prof. Dr. Harald Ritz

Technische Hochschule
Mittelhessen

Fachbereich MNI
Wiesenstraße 14
35390 Gießen
harald.ritz@mni.thm.de

Dr. Holger Hoheisel

CCP Software GmbH

Rudolf-Breitscheid-Straße 1-5
35037 Marburg
hhoheisel@ccpssoft.de

Kategorie

Bachelorarbeit

Schlüsselwörter

Softwarelizenzmanagement, Software Asset Management (SAM), Managed Services, Managed Service Provider (MSP), SWOT-Analyse, Requirements-Engineering, Marktanalyse, Marketing, Kundennutzen.

Zusammenfassung

In den letzten Jahren hat sich Software Asset Management (SAM), oder auch Softwarelizenzmanagement, als wesentlicher Bestandteil eines Unternehmens herausgestellt. Grund für die Umsetzung eines solchen Managements sind unter anderem wirtschaftlich hohe Kosten, bei einer inkorrekten Nutzung von erworbenen Softwarelizenzen. Immer mehr Hersteller verlangen eine compliance-gerechtere Nutzung ihrer Softwarelösungen und wollen dies mit Hilfe sogenannter Software Audits überprüfen. Die Verwaltung aller sich im Unternehmen im Einsatz befindlichen Softwarelizenzen, ist eine anspruchsvolle Aufgabe und verlangt organisierte Prozesse und Strukturen.

Aufgrund der Komplexität fangen Unternehmen an ihr Softwarelizenzmanagement auszulagern, um sich besser auf ihr Kerngeschäft konzentrieren zu können. Ein verbreitetes Konzept zur Unterstützung ist die Inanspruchnahme eines sogenannten Managed Service. Bestehend aus Lizenzmanagern und entsprechender Software sollen Softwarebestände eines Unternehmens erfasst, analysiert, ausgewertet und compliance-konform gemacht werden.

Das Ziel dieser Arbeit besteht darin, die grundlegende Forschungsfrage zu beantworten, wie ein Managed Service dem SAM effektiv unterstützend zur Seite stehen kann. Weiterhin soll in dieser Arbeit ein bereits funktionierender SAM Managed Service eines Unternehmens analysiert und verbessert werden. Als Ergebnis soll ein verbessertes Managed-Service-Konzept entstehen.

Im Blickpunkt der Arbeit steht die CCP Software GmbH mit dem Ziel, ihren SAM Managed Service zu verbessern und insbesondere den Kundennutzen der Leistungen zu erhöhen. Hierzu wird der SAM Managed Service der CCP Software GmbH analysiert und bewertet. Eine SWOT-Analyse gibt aufschlussreiche Informationen über den aktuellen Zustand und zeigt deutliche Verbesserungspotentiale auf. Zudem finden sich ausführliche Gegenüberstellungen der einzelnen Managed Service Provider (MSP) und deren Leistungen hinsichtlich der SAM Managed Services wieder. Somit wird klar, welche Vorteile die unterschiedlichen Wettbewerber mitbringen und welche Defizite sie aufweisen. Mit Hilfe dieser Informationen kann ein verbesserter SAM Managed Service für die CCP Software GmbH erarbeitet werden. Das Interesse der Kunden wird ebenfalls aufgegriffen und es besagt, dass die Übertragung von finanziellen Risiken auf den MSP ein bedeutender Aspekt ist. Zudem sind ein voller Leistungsumfang bei einem minimalen Aufwand und geringen Kosten sowie ein variables Kostenmodell bei der Inanspruchnahme eines Managed Service von hoher Relevanz.

Zielführend sind die Argumente zur verbesserten Vermarktungsstrategie und der Erhöhung des Kundennutzens. Entscheidende Faktoren für die Kundenakquise sind aussagekräftige Leistungsbeschreibungen, vielfältige Supportmöglichkeiten, stark beworbene Zertifizierungsgrade und eine verbesserte interne Kommunikation zwischen Marketing, Vertrieb und den Managed-Service-Lizenzmanagern. Außerdem sollten sich MSP der Bedeutung von IT-Service-management (ITSM) und Enterprise Service Management (ESM) bewusst werden, proaktiv, anpassungsfähig und innovativ agieren und schon früh auf den Trend der Server- und Cloud-Modelle aufsteigen, um in der Digitalisierung präsent zu sein und den Kunden den neuesten Stand anbieten zu können.

Konzeption und Implementierung eines flexiblen Lagerdashboards für SAP EWM mit Microsoft Power BI

Anna Ginovker

Technische Hochschule
Mittelhessen

Fachbereich MND
Wilhelm-Leuschner-Str. 13
61169 Friedberg
E-Mail:
anna.ginovker@mnd.thm.de

Prof. Dr. Harald Ritz

Technische Hochschule
Mittelhessen

Fachbereich MNI
Wiesenstr. 14
35390 Gießen
E-Mail:
harald.ritz@mni.thm.de

Roland Fakesch

Flexus AG

Abtlg. Mobile
John-Skilton-Str. 2
97074 Würzburg
E-Mail:
r.fakesch@flexus.net

Kategorie

Bachelorarbeit

Schlüsselwörter

Visualisierung, Dashboard, Lagerwesen, Kennzahlen, MS Power BI, SAP EWM

Zusammenfassung

Kennzahlen nehmen heute in Unternehmen eine signifikante Rolle ein und dienen oft als Grundlage für Entscheidungen, sowohl strategisch, taktisch und operativ. Hierzu gehören auch die intralogistischen Abläufe, deren ständige Verbesserung zu signifikanten Wettbewerbsvorteilen führen kann. Durch die fortgeschrittene Digitalisierung ist die Erfassung der intralogistischen Prozesse mittlerweile auf einem sehr hohen Niveau. Ermöglicht wird das durch den Einsatz starker ERP-Systeme, wie z.B. SAP EWM. Diese umfangreiche Datenverfügbarkeit bietet großes Potential zur Auswertung, Steuerung und Optimierung der intralogistischen Prozesse.

Die Arbeit hat zum Ziel, diese Datenverfügbarkeit und deren komplexe Interdependenzen mithilfe eines Lagerdashboards zu entzerren und derart zu visualisieren, dass daraus gewonnene Erkenntnisse das Lagerwesen sinnvoll unterstützen. Kern der Aufgabe ist somit die Visualisierung relevanter Daten aus SAP EWM mit MS Power BI und deren Übergabe mittels OData zwischen den beiden Anwendungen.

Hierzu findet in der Arbeit zunächst eine Auseinandersetzung mit den benötigten Grundlagen statt. Dazu gehören wesentliche Prozesse im Lagerwesen, die Systeme SAP EWM und MS Power BI, sowie Nutzen und Funktion von Dashboards.

Darauf aufbauend werden Kundenanforderungen definiert, und mithilfe einer lagerprozessbasierten Analyse die Funktionsanforderungen an das Lagerdashboard erarbeitet. Daraus werden die zu

visualisierenden Kennzahlen, Daten und Informationen abgeleitet und abschließend die so erarbeiteten Anforderungen in einem Pflichtenheft zusammengefasst.

Weiterhin findet eine Konzeption als Basis für die anstehende Umsetzung statt. Zunächst werden dafür die vorhandenen Daten in SAP EWM gesichtet und anschließend die technische und visuelle Konzeption des Lagerdashboards vorgenommen. Die eigentliche Umsetzung beginnt mit der technischen Verknüpfung zwischen den beiden Systemen SAP EWM und MS Power BI mittels OData und der visuellen Umsetzung des Lagerdashboards. Die so in Power BI transportierten Datentabellen werden angepasst, ergänzt und logisch miteinander verknüpft, so dass ein komplexes Datenmodell als Grundlage für alle erarbeiteten Dashboard-Varianten entsteht. Funktionstests und Verbesserungen des erstellten Prototyps schließen die Arbeit ab.

Beispielhaft für ein realisiertes Lagerdashboard ist das Monitoring der Lagerauslastung. Die darin angezeigten Informationen und Kennzahlen zeigen dem Benutzer primär die aktuelle Belegung vorhandener Lagerplätze im Lager als absolute Kennzahlen nach Gewicht, Volumen, Anzahl HU's usw. an. Flankiert werden diese durch Balkendiagramme mit der Darstellung der Lagertypen und -Bereiche. Das Anklicken einzelner Balken führt zur Filterung und Anzeige der entsprechenden Kennzahlen des so ausgewählten Bereiches und zur Darstellung der Auslastung.

Zusammenfassend ist es anhand der erarbeiteten Anforderungen an das Lagerdashboard gelungen, relevante Daten in SAP EWM zu identifizieren, in einem Datenmodell aufzubereiten, um eigene Ausarbeitungen zu ergänzen, und nach MS Power BI zu transportieren. Es entstanden insgesamt 14 ansprechend visualisierte Lagerdashboard-Varianten mit über 100 sog. *Visuals*. Das Ergebnis zeigt dabei funktionsorientierte Lagerdashboards mit unterschiedlichen Schwerpunkten in Bezug auf die jeweiligen Lagerprozesse und -zustände.

Die im Rahmen dieser Arbeit erarbeiteten Lagerdashboards und vor allem deren Datengrundlage basieren auf einem unternehmensinternen Testsystem der FLEXUS AG¹. Es wurde deshalb bei der Umsetzung auf starre Lösungen verzichtet, so dass bei einer anstehenden Implementierung im Kundenumfeld Lagerdashboards flexibel sowohl im Design als auch in der Darstellung und Mix der Kennzahlen mit minimalem Aufwand angepasst werden können.

Das erzielte Ergebnis hat das Potential die angestrebte Unterstützung für das operative Lagerwesen ergänzend zum SAP EWM zu leisten.

¹ Softwareentwickler (Beratung, Programmierung, Implementierung) im Bereich Intralogistik, vorzugsweise auf Basis SAP.

Implementierung eines Visualisierungskonzepts für die grafische Darstellung von Daten und Informationen innerhalb des Controllings

Lisa Hausner

Technische Hochschule
Mittelhessen

Fachbereich MND
Wilhelm-Leuschner-Str. 13
61169 Friedberg
lisa.hausner@mnd.thmd.de

Prof. Dr. Harald Ritz

Technische Hochschule
Mittelhessen

Fachbereich MNI
Wiesenstraße 14
35390 Gießen
harald.ritz@mni.thm.de

Manuel Sawusch

Bosch Rexroth AG

DC-IH/CTG
Zum Eisengießer 1
97816 Lohr am Main
manuel.sawusch@boschrex-roth.de

Kategorie

Bachelorarbeit

Schlüsselwörter

Visualisierungen, Informationsdesign, Business Intelligence, Dashboards, Gestaltungsrichtlinien, IBCS, Microsoft Power BI, Tableau

Zusammenfassung

In jedem Unternehmen werden Daten produziert und genutzt. Speziell im Controlling werden durch Daten und Informationen wichtige strategische, operative und taktische Entscheidungen getroffen, die weitreichende Folgen haben können. Umso wichtiger ist es, dass die Daten und die daraus abgeleiteten Informationen richtig verarbeitet, verstanden und kommuniziert werden. Visualisierungen können dabei helfen, aber auch verwirren, wenn sie unsachgemäß eingesetzt werden. Deshalb ist es sinnvoll, sich bei der Verwendung von Visualisierungen an Gestaltungsrichtlinien zu halten. Dadurch wird die Erstellung und das Verständnis beschleunigt und vereinfacht.

Das Ziel der Bachelorarbeit ist es, einen Leitfaden für die Erstellung von Visualisierungen zu verfassen. Dazu werden Empfehlungen aus der Literatur herangezogen, wie z.B. die International Business Communication Standards der IBCS Association oder die Empfehlungen des Informationsdesigners Stephen Few. Die Umsetzung des Leitfadens wird anhand von zwei Softwarebeispielen, Microsoft Power BI und Tableau, überprüft. Darüber hinaus wird ein Dashboard für die Fixkostenanalyse für das Controlling der Business Unit Industriehydraulik der Bosch Rexroth AG implementiert, das ebenfalls den Regeln des erstellten Leitfadens folgen soll.

Die Ergebnisse zeigen, dass die Empfehlungen der Literatur zur Darstellung von Visualisierungen sehr ähnlich

sind und die Erstellung eines Leitfadens gut möglich ist. Allerdings muss man sich bei der Erstellung fragen, wie weitreichend und detailliert die Regeln sein sollten. Darüber hinaus ist es für Unternehmen ratsam, ein eigenes Regelwerk zu erstellen, das auf die eigenen Bedürfnisse und Wünsche zugeschnitten ist. Dieses Regelwerk sollte dann unternehmensweit gelten und angewendet werden.

Die Umsetzung des Leitfadens mit der Beispielsoftware Microsoft Power BI und Tableau zeigt, dass nicht alle Empfehlungen eins zu eins umgesetzt werden können. Dies liegt vor allem an den vorgegebenen Individualisierungsmöglichkeiten der Visualisierungssoftware. Erst mit ergänzenden Erweiterungen, die innerhalb der Software heruntergeladen werden können, lassen sich die Empfehlungen vollständig umsetzen. Die Erweiterungen, wie z.B. das Power BI zertifizierte Add-on Zebra BI oder die graphomate extensions, setzen die IBCS-Regeln nahezu perfekt um und können zudem nach eigenen Wünschen angepasst werden. Vor der Verwendung solcher Erweiterungen muss das Unternehmen jedoch prüfen, ob der Einsatz unter Sicherheitsaspekten zulässig ist. Weiterhin sind nicht alle Funktionen für jede Visualisierungssoftware kostenfrei verfügbar.

Die Umsetzung im Unternehmen zeigt aber auch, dass es individuelle Wünsche gibt, welche nicht immer mit den Empfehlungen übereinstimmen. Wenn Dashboards zum ersten Mal eingeführt werden, brauchen auch die Mitarbeiter Zeit, um sich darauf einzustellen. Wie bereits erwähnt, ist es für Unternehmen ratsam ein eigenes individuelles und unternehmensweites Regelwerk zu erstellen. Um die Akzeptanz zu erhöhen, ist es empfehlenswert das Projekt gemeinsam in Workshops zu bearbeiten. Vor der Erstellung eines eigenen Leitfadens sollten die Empfehlungen auch mit den im Unternehmen vorhandenen Softwaremöglichkeiten abgeglichen werden, um sicherzustellen, dass alle Ratschläge auch in der Praxis umgesetzt

werden können. Darüber hinaus sollte der Leitfaden immer auf dem neusten Stand gehalten werden, da z.B. Softwareupdates zu Änderungen führen können.

Es sei darauf hingewiesen, dass es sich bei dem Leitfaden schlussendlich nur um Empfehlungen handelt, die nicht strikt durchgesetzt werden müssen, sondern einen Gesamteindruck von guter Visualisierung vermitteln sollen.

Literatur

Few, Stephen: Information Dashboard Design – Displaying Data for at-a-Glance Monitoring, 2. Aufl., Burlington: Analytics Press, 2013.

IBCS Association: International Business Communication Standards, IBCS® Version 1.2, Hilden: IBCS Media, 2022.

Kohlhammer, Jörn; Wiener, Andreas; Proff, Dirk U: Visual Business Analytics – Effektiver Zugang zu Daten und Informationen, 2. Aufl., Heidelberg, München: Dpunkt.Verlag, 2018.

Vergleichende Bewertung von Data-Warehouse-Landschaften in der Cloud am Beispiel von Microsoft Azure Synapse Analytics und Snowflake

Thao-An Hua

Technische Hochschule Mittelhessen

Fachbereich MND
Wilhelm-Leuschner-Str. 13
61169 Friedberg
thao-an.hua@mnd.thm.de

Prof. Dr. Harald Ritz

Technische Hochschule Mittelhessen

Fachbereich MNI
Wiesenstraße 14
35390 Gießen
harald.ritz@mni.thm.de

Wadim Myza (B.Sc.)

INFOMOTION GmbH

BU Veltzé
Westhafenplatz 1
60327 Frankfurt am Main
wadim.myza@infomotion.de

Kategorie

Bachelorarbeit

Schlüsselwörter

Business Intelligence, Cloud-Computing, Softwareevaluation, Bewertungsmatrix, Microsoft Azure Synapse Analytics, Snowflake

Zusammenfassung

Die Welt befindet sich in einem zunehmend rapiden gesellschaftlichen Wandel. Vieles soll schneller, virtueller und digitaler werden. Infolge der digitalen Transformation ist das unverzügliche Treffen fundierter Entscheidungen von signifikanter Bedeutung, um kompetitiv zu bleiben. Durch die frühzeitige Reaktion auf Änderungen im Marktumfeld kann ein Wettbewerbsvorteil gegenüber der Konkurrenz erzielt werden. Zur Entscheidungsunterstützung greifen viele Unternehmen auf Data-Warehouse-Landschaften in der Cloud zurück, um große Datenmengen zu analysieren. Die zentralen Merkmale des Cloud-Computings, wie die Skalierbarkeit, tragen zu einer effizienten Entscheidungsfindung bei.

Die adäquate Auswahl einer Cloud-Datenplattform-Lösung stellt eine Herausforderung für Unternehmen dar. Angesichts der stetig voranschreitenden Digitalisierung und jene damit einhergehende unterschiedliche Reifegrade der am Markt verfügbaren Auswahlmöglichkeiten ergibt sich für viele Unternehmen die Frage, welche Cloud-Datenplattform und dessen Datenlösung die beste Unterstützung bietet. Infolgedessen ist das Untersuchungsziel der Bachelorarbeit die Konzeption einer Bewertungsmatrix zur fundierten Auswahl einer Cloud-Datenplattform-Lösung. Am Beispiel der Gegenüberstellung des PaaS-Produkts Azure Synapse Analytics von Microsoft (MS) und des SaaS-Produkts Snowflake ist zu eruieren, welche Lösung in Bezug auf die beleuchteten Anforderungen am besten geeignet ist.

Mit Microsoft steht Unternehmen einer der größten Cloud-Anbieter zur Verfügung. In Anbetracht der breiten

Palette an Produkten, wie beispielsweise Synapse Analytics, werden allgemein moderne Lösungen für diverse Anwendungsszenarien angeboten. Snowflake ist eine vom gleichnamigen Unternehmen entwickelte SaaS-Lösung. Der Komplettservice bietet eine vollständige Verwaltung der Infrastruktur in der Cloud an, sodass der Fokus auf die Analyse und Verwaltung der Daten gelenkt werden kann.

Zur Ermittlung relevanter Kriterien wurden verschiedene Quellen gesichtet. Die gewonnenen Erkenntnisse sind in die Konzeption der Bewertungsmatrix eingeflossen, die aus 17 Kriterien besteht. Diese sind in funktionale und qualitative Anforderungen gegliedert und in den sieben Kategorien Administration, Architektur, Interoperabilität, Kosten, qualitative Kriterien, Data Governance und Flexibilität unterteilt.

Des Weiteren wurde jeweils ein Datenmodell in MS Azure Synapse Analytics und Snowflake, im Kontext eines fiktiven Anwendungsszenarios umgesetzt. Die vierstufige Data-Warehouse-Grundarchitektur stellte hierbei die Basis. Als Datenquelle wurde ein Ausschnitt der Daten der aktuellen COVID-19-Pandemie verwendet. Bei der praktischen Umsetzung stellte sich dabei heraus, dass die Implantierung je nach gesammelten Erfahrungen ein mehr oder weniger kompliziertes Unterfangen darstellt. Auf Basis der jeweiligen Datenmodelle wurde jeweils ein Dashboard mittels Power BI Desktop entworfen und realisiert. Anhand dessen und durch die jeweilige Umsetzung konnten sowohl für MS Azure Synapse Analytics als auch für Snowflake praktische Beispiele zur Bewertung der einzelnen Kriterien herangezogen werden. In diesem Zusammenhang wurden beispielsweise die jeweiligen Kosten für den Entwicklungszeitraum berechnet und Performance-Tests durchgeführt.

Im Hinblick auf die beleuchteten Kriterien gelten sowohl MS Azure Synapse Analytics als auch Snowflake als moderne Analyseplattform. Bei der vergleichenden Bewertung hat sich zum Betrachtungszeitpunkt herausgestellt, dass Snowflake in Bezug auf die Anforderungen an eine

moderne Analyseplattform die sieben Kriterien Auto-Resume, Bereitstellung, Skalierbarkeit, Performance, Einsatz von Drittanbieter-Tools, Abrechnung und Wartbarkeit besser qualifiziert erfüllt als MS Azure Synapse Analytics. Trotzdem kann auch MS Azure Synapse als moderne Analyseplattform einen Mehrwert für das Unternehmen liefern, da die individuellen Anforderungen eines Unternehmens je nach spezifischen Anwendungsfall variieren.

Literatur

Hartmut, B.; Hartmut, H.: Vertrieb und Marketing in der digitalen Welt, Springer Gabler, Wiesbaden, 2018

Kraynak, J.; Baum, D.: Cloud Datawarehousing, 2. Auflage, John Wiley & Sons, New York, 2020

Schnider, D.; Jordan, C.; Welker, P.; Wehner, J.: Data Warehouse Blueprints, Carl Hanser Verlag GmbH & Co. KG, München, 2016

Norman, B.: Treiber einer „Data Warehouse as a Service“-Lösung, In: Finger, R. (Hrsg.): *BI & Analytics in der Cloud: Architektur, Vorgehen und Praxis*, 1. Auflage, dpunkt.verlag, Heidelberg, 2018

Design und Implementierung einer modernen „State of the Art“ SAP S/4HANA-Cloud-Anwendung mit einer Benutzeroberfläche basierend auf SAP Fiori

Christoph Günter Rudewig

Technische Hochschule
Mittelhessen

Fachbereich MND
Wilhelm-Leuschner-Str. 13
61169 Friedberg

E-Mail:

christoph.guenter.rudewig@mnd.thm.de

Prof. Dr. Harald Ritz

Technische Hochschule
Mittelhessen

Fachbereich MNI
Wiesenstraße 14
35390 Gießen

E-Mail: harald.ritz@mni.thm.de

Prof. Dr. Frank Kammer

Technische Hochschule
Mittelhessen

Fachbereich MNI
Wiesenstrasse 14
35390 Gießen

E-Mail:

frank.kammer@mni.thm.de

Kategorie

Bachelorarbeit

Schlüsselwörter

SAP S/4HANA, SAP S/4HANA Cloud, ABAP-RESTful-Programmiermodell, ABAP Core Data Services Views, SAP Fiori, SAP Fiori for Elements, SAP Business Technology Platform, Side-By-Side-Erweiterungen

Zusammenfassung

Lange Zeit war es möglich innerhalb des ERP-Systems von SAP Erweiterungen vorzunehmen, um das ERP-System von SAP in gewissen Bereichen auf die Bedürfnisse des jeweiligen Unternehmens zuzuschneiden. Mit der Einführung von SAP S/4HANA im Jahre 2015 bietet SAP ein neues ERP-System an. Seit dem Jahr 2016 stellt SAP zudem mit SAP S/4HANA Cloud ein komplettes ERP-System in der „Cloud“ zur Verfügung. Vor diesem Hintergrund wird im Rahmen der Bachelorarbeit erörtert, ob es mit SAP S/4HANA Cloud weiterhin möglich ist Erweiterungen vorzunehmen.

Die Untersuchung zeigte, dass Erweiterungen mit SAP S/4HANA Cloud im traditionellen Sinne, wie man sie bisher von On-Premises-ERP-Systemen von SAP kannte, nicht mehr möglich sind. Das liegt daran, dass der „digitale Kern“ von SAP S/4HANA Cloud, in dem die Standardprozesse von SAP S/4HANA Cloud implementiert sind, nicht geändert werden darf. Um Erweiterungen an SAP S/4HANA Cloud vorzunehmen gibt es die Möglichkeiten der In-App-Erweiterungen und der Side-By-Side-Erweiterungen. Die In-App-Erweiterungen sind jedoch sehr begrenzt und speziell für Fachanwender gedacht. Um eine eigene Cloud-Anwendung zu entwickeln, muss sich der Side-By-Side-Erweiterungen bedient werden.

Side-By-Side-Erweiterungen werden auf der SAP Business Technology Platform (BTP) entwickelt. Die

SAP BTP existiert unter diesem Namen seit Anfang des Jahres 2021. Es handelt sich um eine Entwicklungs- und Verwaltungsplattform, mit deren Funktionen es möglich ist, Cloud-Anwendungen zu entwickeln, zu betreiben und in SAP S/4HANA-Systeme zu integrieren.

In diesem Zusammenhang stellt die SAP BTP die ABAP-Environment als Entwicklungsumgebung zur Verfügung, um auf Basis des ABAP-RESTful-Programmiermodells Cloud-Anwendungen zu entwickeln. Das ABAP-RESTful-Programmiermodell basiert u.a. auf neuen SAP-Technologien wie ABAP Core Data Services Views und SAP Fiori. Mit ABAP Core Data Services Views wird eine neue Datenschicht zwischen den Datenbanktabellen und der Anwendungslogik eingeführt, da auf die Datenbanktabellen der SAP-HANA-Datenbank in SAP S/4HANA-Systemen nicht mehr direkt zugegriffen werden soll. Mit SAP Fiori wurde eine neue Benutzeroberfläche eingeführt, die den Anspruch hat eine benutzerfreundliche Bedienung von Anwendungen zu ermöglichen.

Auf Basis dieser Erkenntnisse wurde eine Cloud-Anwendung entworfen und anschließend auf der SAP Business Technology Platform implementiert. Die SAP-Fiori-Oberfläche wurde dabei mit dem Framework SAP Fiori Elements programmiert. Dabei stellte sich heraus, dass es möglich ist ohne Kenntnisse der Programmiersprache ABAP oder dem Framework SAPUI5 eine Cloud-Anwendung mit Grundfunktionen wie dem Anlegen, Ändern, Lesen und Löschen von Daten und einer Benutzeroberfläche basierend auf SAP Fiori zu entwickeln.

Literatur

Colle, R.; Dentzer, R.; Hrastnik, J.: Core Data Services für ABAP, 2. Auflage, Rheinwerk, Bonn, 2020

Tiebold, A.; Künold, B.; Jungk, P.: ABAP-Entwicklung auf der SAP Cloud Platform, Rheinwerk, Bonn, 2020

Automatisierung des Aufbaus flexibler Infrastrukturen analytischer Systeme in Cloud-Umgebungen

Viktor Schneider

Technische Hochschule
Mittelhesen

Fachbereich MND
Wilhelm-Leuschner-Str. 13
61169 Friedberg
viktor.schneider@mnd.thm.de

Prof. Dr. Harald Ritz

Technische Hochschule
Mittelhesen

Fachbereich MNI
Wiesenstr. 14
35390 Gießen
harald.ritz@mni.thm.de

Sebastian Becker

INFOMOTION GmbH

BU Duheric
Westhafenplatz 1
60327 Frankfurt am Main
sebastian.becker@infomotion.de

Kategorie

Masterarbeit

Schlüsselwörter

Amazon Web Services (AWS), Analytische Systeme, Big Data, Cloud Computing, Data Lake, Data Warehouse (DWH), Infrastructure-as-Code (IaC), Konfigurationsmanagement, Microsoft Azure, Provisionierung.

Zusammenfassung

Die in den Unternehmen vorkommenden Anwendungsfälle von Datenauswertungen weisen eine große Vielfalt auf. Als Konsequenz wachsen kontinuierlich die Anzahl und die Funktionsvariabilität von Softwaresystemen, die diese Aufgaben unterstützen. Der gemeinsame Einsatz dieser Softwareteile führt zu komplexen und verschiedenartigen Architekturen, deren Realisierung einen immensen Aufwand erfordert. Außer klassischen Data-Warehouse-Systemen (DWH) finden sich seit einiger Zeit auch Big-Data-Anwendungen in der analysebezogenen IT-Landschaft von Unternehmen. Die Vielfalt und Komplexität dieser Architekturen behindert, verlangsamt und verteuert die Erprobung neuer und innovativer Zusammensetzungen aus analytischen Systemen im Rahmen von Proof-of-Concept-Studien.

Cloud-Computing-Dienste vereinfachen die Bereitstellung solcher Systemverbünde durch die Automatisierung der Ressourcenbereitstellung und -betriebs. Der enorme Zeitaufwand, die Fehleranfälligkeit, die evtl. zu einem mühsamen Neuaufbau des ganzen Systemverbunds führen kann, sowie die eingeschränkte Wiederverwendbarkeit eines manuell durchgeführten Aufbaus lassen sich dadurch aber nicht vermeiden. Eine Alternative zur manuellen Vorgehensweise bietet der Infrastructure-as-Code-Ansatz (IaC). Die Bereitstellung von Softwaresystemen und der darunter liegenden Infrastruktur lässt sich damit als Code einer universellen Programmiersprache oder einer domänenspezifischen Sprache definieren. Dies ermöglicht die Verwendung von Methoden aus dem Software Engineering, wie die Versionsverwaltung von

Code, den Einsatz wiederverwend- und parametrisierbarer Komponenten, die Definition und Durchführung von Tests u.v.a.m.

Im Rahmen dieser Arbeit ist es trotz des breiten Spektrums an Architekturvarianten analytischer Systeme gelungen, mit Hilfe des IaC-Ansatzes ein gesamtheitliches Konzept für die Aufbauautomatisierung in mehreren Cloud-Umgebungen zu entwickeln und seine Realisierbarkeit anhand eines Prototyps nachzuweisen. Sowohl die klassischen DWH als auch die unterschiedlichen Arten von Big-Data-Anwendungen lassen sich zu einer sie umfassenden mehrschichtigen Referenzarchitektur vereinen. Die Hauptbestandteile dieser Architektur sind ein Data Lake für die Rohdaten, SQL- und NoSQL-DBMS, ETL-Werkzeuge, Big-Data-Processing-Engines sowie Datenanalyse- und Visualisierungswerkzeuge.

Alle Systeme aus dieser Referenzarchitektur können auf virtuellen Maschinen eigenadministriert, in Form einer SaaS-Lösung oder eines PaaS-Dienstes betrieben werden. Obwohl sich die Cloud-Umgebungen in ihrem Angebot von Diensten unterscheiden, ist ihre Architektur ähnlich gestaltet. Die in dieser Arbeit analysierten Cloud-Umgebungen Amazon Web Services (AWS) und Microsoft Azure unterscheiden sich in Bezug auf IaaS-Dienste nur marginal. Aber auch im Fall von Big-Data-Diensten greifen die beiden Cloudanbieter auf dieselben Processing Engines und Machine-Learning-Frameworks zurück. Insgesamt entstand der Eindruck, dass beide Cloudanbieter den Fokus mehr auf Big-Data-Technologien als auf komplette DWH-Lösungen legen.

Jedes Softwaresystem kann – unabhängig vom Cloud-dienst, auf dem es läuft – von diesem Automatisierungskonzept erfasst werden, solange es die Kriterien einer Ressource erfüllt. Eine Ressource hat zugreifbare Eigenschaften, einen überwachbaren Zustand und kann erzeugt, modifiziert und zerstört werden. Ressourcen lassen sich zu wiederverwendbaren Modulen zusammensetzen. Durch die Parametrisierung von Modulen können die Softwaresysteme schnell und flexibel zu vielfältigen Systemverbünden kombiniert werden. Für die Erfassung

von neuen Systemen in die Automatisierungslösung wurde außerdem ein Prozessmodell erstellt.

Die Implementierung des Konzepts erfolgt unter Einsatz von speziellen IaC-Werkzeugen. Sie verwenden die Programmierschnittstellen von Clouddiensten oder den Softwaresystemen selbst, um die Systeme in der Cloud-Umgebung zu provisionieren und zu konfigurieren. Im Rahmen einer Vorauswahl wurden aus im Internet aufgefundenen IaC-Werkzeugen vier Instrumente mit dem nötigen Funktionsumfang bestimmt. Für die Endauswahl fand eine Evaluierung anhand von in der Literatur ermittelten und anforderungsgerechten Kriterien statt.

Für den Prototyp wurde ein DWH-System bestehend aus einem Data Lake, einem RDBMS, einem ETL-Werkzeug sowie einer Analyse- und Visualisierungsanwendung entwickelt. Die drei Clouddiensten IaaS, PaaS und SaaS sind jeweils von mindestens einem DWH-Systemteil repräsentiert. Für jedes Systemteil wurden Ressourcen und Module sowohl für die AWS- als auch die Azure-Cloud erstellt, erfolgreich getestet und damit die cloudübergreifende Eigenschaft des Konzepts nachgewiesen. Weiterhin wurde in der AWS-Cloud eine RDBMS-Ressource auf einer virtuellen Maschine durch eine AWS-PaaS-Ressource ersetzt, um die flexible Anwendung der Lösung zu zeigen.

Literatur

Baars, Henning; Kemper, Hans-Georg: Business Intelligence & Analytics – Grundlagen und praktischen Anwendungen. 4. überarb. und erw. Aufl., Wiesbaden: Springer Vieweg Verl., 2021.

Bauer, Andreas; Günzel, Holger (Hrsg.): Data-Warehouse-Systeme – Architektur Entwicklung Anwendung. 4. überarb. und erw. Aufl., Heidelberg: dpunkt.verlag, 2013.

Brikman, Yevgeniy: Terraform: Up & Running. 2. Aufl., Sebastopol: O'Reilly Media, Inc., 2019.

Erl, Thomas; Mahmood, Zaigham; Puttini, Ricardo; Wise-Martinez, Pamela J.: Cloud computing – Concepts, technology & architecture, Upper Saddle River usw.: Prentice-Hall/Pearson, 2013.

Freiknecht, Jonas; Papp, Stefan: Big Data in der Praxis – Lösungen mit Hadoop, Spark, HBase und Hive – Daten speichern, aufbereiten, visualisieren. 2. erw. Aufl., München: Carl Hanser Verl., 2018.

Gorelik, Alex: The Enterprise Big Data Lake – Delivering the Promise of Big Data and Data Science, Sebastopol usw.: O'Reilly Media, Inc. Verl., 2019.

Heap, Michael: Ansible - From Beginner to Pro, New York: Apress Verlag, 2016.

Morris, Kief: Infrastructure as code – Managing servers in the cloud, Sebastopol: O'Reilly Media, Inc., 2016.

o.V.: Übersicht über Amazon Web Services – Whitepaper zu AWS, o.O., 2021, online im Internet: URL: https://docs.aws.amazon.com/de_de/whitepapers/last-test/aws-overview/aws-overview.pdf. (Abruf: 08.03.22)

o.V.: Übersicht über die Dienste der Azure-Plattform, o.O., 2022, online im Internet: URL: <https://docs.microsoft.com/de-de/azure/?product=all>. (Abruf: 08.03.22)

Savill, John: Mastering Microsoft Azure Infrastructure Services, Indianapolis: Syber Verl., 2015.

Wittig, Michael; Wittig, Andreas: Amazon Web Services in Action, New York: Manning Publications Co. LLC, 2018.

Standardisierung von automatisierten Bestellprozessen mit Peppol am Beispiel der Friedhelm Loh Group

Max Münker

Technische Hochschule
Mittelhessen

Fachbereich Mathematik,
Naturwissenschaften und
Datenverarbeitung
Wilhelm-Leuschner-Str. 13
61169 Friedberg
max.muenker@mnd.thm.de

Prof. Dr. Armin
Wagenknecht

Technische Hochschule
Mittelhessen

Fachbereich Mathematik,
Naturwissenschaften und
Informaik
Wiesenstraße 14
35396 Gießen
armin.wagen-
knecht@mni.thm.de

Kategorie

Masterarbeit

Schlüsselwörter

Peppol, Electronic Data Interchange, EDI,
Standardisierung, TOGAF

Zusammenfassung

Electronic Data Interchange (EDI) hat sich seit der Entwicklung in den 1970er Jahren als Technologie zum optimierten Austausch von Geschäftsdaten im B2B-Bereich etabliert. Die Automatisierung von Bestellprozessen mit EDI bietet Unternehmen verbesserte Prozesse mit der Vermeidung von fehleranfälligen manuellen Schritten. Trotz der erheblichen Vorteile von EDI ist die damit verbundene Integration der Geschäftspartner aufgrund der unzureichenden Standardisierung mit einem hohen Kosten- und Zeitaufwand verbunden. Aus der Verwendung aktuell verbreiteter EDI-Formate, wie ANSI X12 oder UN/EDIFACT, resultiert ein erhöhter Implementationsaufwand je anzubindendem Geschäftspartner.

Die im Jahr 2008 bis 2012 entwickelte Technologie Pan-European Public Procurement Online (Peppol) bietet eine Alternative zu herkömmlichen Standards. Es stellt einen Zusammenschluss aus Netzwerkteilnehmern zum einheitlichen Austausch von Geschäftsdokumenten dar. Neben der in einer EU-Reform vorgeschriebenen Nutzung von Peppol zum Versand elektronischer Rechnungen an öffentliche Stellen, unterstützt das Peppol-Netzwerk auch den Austausch von weiteren Dokumenten, wie Bestellungen. Nach der Einführung in Teilen Europas ist Peppol nun bereits in 37 verschiedenen Ländern im Einsatz, inklusive Teilen Asiens, Neuseeland, Australien und der USA. Im August 2021 sind bereits ca. 2,5 Millionen Unternehmen als Netzwerkteilnehmer bei Peppol registriert, wovon ca. 81.800 Unternehmen den Austausch von Bestellungen unterstützen.

Die Vorteile einer Standardisierung von EDI-Bestellprozessen mithilfe von Peppol wurden im Rahmen dieser Masterarbeit in Kooperation mit der international agierenden Friedhelm Loh Group (kurz F.L.G. u.a. Muttergesellschaft der Rittal GmbH & Co. KG) analysiert.

Zur strukturierten Erarbeitung des Standardisierungsvorhabens wurde mit Hilfe des TOGAF-Frameworks die zugrundeliegende IT-Architektur im Bereich EDI mit Bezug auf die F.L.G. betrachtet. Die Gegenüberstellung der aktuellen zur zukünftig angestrebten Architektur ermöglicht die Erstellung einer langfristigen Planung zur Standardisierung von EDI. Eine strategische Betrachtung von EDI anhand der IT-Architektur ermöglicht eine zielgerichtete Entscheidungsfindung für die Geschäftsführung des Unternehmens. Neben der technischen Modernisierung wurden die kaufmännischen Vorteile einer Standardisierung mit Peppol verdeutlicht. Als Mittel zur Beurteilung wurden exemplarisch Umsatz- und Kostenanalysen des F.L.G. Geschäftsbereiches Rittal International herangezogen.

Zur Etablierung von Peppol als EDI-Standard wird der F.L.G. im Rahmen der IT-Architekturbetrachtung ein Fünfjahresplan vorgeschlagen. Neben der initialen Integration der bestehenden Informationssysteme an das Peppol-Netzwerk, wird die internationale Einführung in mehrere Phasen und unterschiedliche Ländergruppen eingeteilt.

Die Masterarbeit zeigt, dass unter der Verwendung von Peppol eine schnellere und kostengünstigere Integration von Geschäftspartnern erreicht werden kann. Nach initialer Einführung wird eine Zeit- und Kostenersparnis von über 95 Prozent im Vergleich zu klassischen EDI-Standards erwartet. Trotz des starken Bezugs zur F.L.G. können die erarbeiteten Ergebnisse auf andere Unternehmen übertragen werden, sodass eine globale Standardisierung grundsätzlich möglich ist.