

The Future of Big Data and Business Operations in the Context of Relational In-Memory Databases

Scenarios for the Year 2025

Master Thesis

Julia Reichelt

Student ID University of Twente: s1616633

Student ID Technical University of Berlin: 354767

To obtain a:

Master of Science (M.Sc.) in Business Administration
(University of Twente)

Master of Science (M.Sc.) in Innovation Management & Entrepreneurship
(Technical University of Berlin)

Submitted to

Prof. Dr. Katharina Hölzle
University of Potsdam

Supported by Victoria Götz, M.Sc.

Dr. Michel Ehrenhard
University of Twente



UNIVERSITY OF TWENTE.



Table of Contents

Management Summary	III
List of Figures	V
List of Tables	V
List of Abbreviations.....	VI
1 Motivation, research question and thesis outline	1
1.1 Problem statement: Real-time decisions to cope with change in the Industry 4.0.....	1
1.2 Thesis goal and research question: RIMDB future developments and Industry 4.0	3
1.3 Laying the foundations: Relevant literature and the research design	3
1.4 Theoretical & practical contribution	4
1.5 Thesis outline	5
2 Literature review	6
2.1. Technology management.....	6
2.1.1 Introduction to competitive strategy.....	6
2.1.2 Competitive strategy and the firm's environment	7
2.1.3 The role of IT in competitive strategy and the value chain	8
2.1.4 From technology strategy to technology foresight	10
2.2 Business operations, the Internet of Things and Industry 4.0	13
2.2.1 Overview over business operations	13
2.2.2 The Internet of Things and value creation	14
2.2.3 Industry 4.0 or the Industrial Internet of Things.....	16
2.3 Formulation of the research question	18
3 Research design and methodology	19
3.1 Data collection procedure	19
3.1.1 Combination approach of systematic and narrative literature review	19
3.1.2 Semi-structured expert interviews and the expert sample	21
3.2 The scenario method.....	25
3.2.1 Definition of key terms and conceptualization	25
3.2.2 The scenario approach based on Reibnitz.....	26
3.2.3 A critical view on the scenario method	29
4 Results	32
4.1 Scenario preparation: Technology Exploration Report (TER).....	32
4.1.1 Relational In-Memory Database Systems (RIMDB): Technology description and research status.....	32

4.1.2 Technology life cycle of RIMDB	34
4.1.3 The active players and the competitive field	36
4.1.4 Brief assessment of the business potential of RIMDB	39
4.1.5 The driving forces for the future of RIMDB	40
4.2 Scenarios.....	51
4.2.1 Scenario 1: The open-source revolution.....	52
4.2.2 Scenario 2: The big data golden age	54
4.4.3 Scenario 3: Platforms for the win.....	56
5 Conclusions for RIMDB and Industry 4.0	58
5.1 Developments relevant for RIMDB.....	58
5.2 Conclusions for Industry 4.0	60
6 Discussion of limitations and future directions.....	63
7 Acknowledgements.....	66
Appendix.....	67
References.....	VII

Management Summary

Decision-makers are faced with a quickening pace of change today, which increases the uncertainty businesses have to deal with. For this purpose, they need to be able to cope with changing environments faster than ever. Therefore, real-time strategic decision-making is gaining in relevance.

In parallel to these developments, the Internet of Things (IoT) is a recent technological trend which has received much attention. The IoT can be described as ubiquitous connectivity of smart devices or objects. These smart connected devices (SCD) are equipped with sensors, and serve to collect data from the environment. Extensive growth is expected for the IoT: By 2020, estimates project between 25 and 50 billion of such SCD as part of the IoT. Considering the enormous expected amount of SCD, the ensuing data volumes that will be produced will also be very large. However, the full potential of these large amounts of data, also called big data, can only be reached when the collected data are analysed. Thus, in order to create value in the IoT, one essential aspect is data analysis. Of the data collected today, only a small fraction is being analysed.

The IoT creates valuable data that can be used for real-time strategic decisions through data analysis. A crucial technology for real-time big data analysis is relational in-memory databases (RIMDB). Examples for RIMDB are SAP HANA, Oracle TimesTen, Microsoft Hekaton or IMDB DB2 with Blu.

Closely related to the IoT is another trend on which governments, researchers and practitioners put a particular focus at the moment: Industry 4.0. The latter can be summarised as the implementation of the IoT in industrial settings through the real-time use of data from various SCD, thereby changing the activities and processes concerned with producing goods or services from resources (i.e. the operations of a business). The potential impact of the IoT for such industrial settings is expected to be particularly extensive, both for individual firms and for the economy in general. The relevance of real-time strategic decisions, enabled by RIMDB, is thus particularly high for the Industry 4.0.

A thorough understanding of RIMDB can guide decision-makers to making more informed decisions when it comes to the consideration and implementation of the big data analysis technology RIMDB.

Similar to all other domains, RIMDB are also subject to change due to various influencing factors and driving forces. Hence, it is useful to not only look at the technology now, but also its future development and how this development can be shaped. In addition to the players involved in the Industry 4.0, RIMDB vendors can gain from this, as they themselves can exert an influence on the future development of the technology.

To provide the strategic guidance to RIMDB vendors as well as players in the Industry 4.0, this thesis answers the following research question: *How can relational in-memory database (RIMDB) developments by 2025 affect Industry 4.0?* The research question consists of two sub-questions: (1) *What are future developments of RIMDB by 2025?* and (2) *What do these future developments mean for Industry 4.0?*

To answer the research question, a technology foresight study was conducted, using the scenario method mainly based on Reibnitz (1988). The purpose of the scenarios is not to predict the future, but to open the space of considered possibilities in decision-making by providing various alternative, possible futures. Supporting the scenario approach, ten experts from research and practice were consulted in thirteen semi-structured, qualitative interviews between November 2015 and March 2016. The software *Parmenides EIDOS™* supported the scenario approach.

The scenarios built for this thesis are based on prior identification of 31 driving forces that can shape the development of RIMDB. Based on further assessment of the influences between those 31 drivers, six key drivers were identified: (1) Demand of real-time response & service acceleration, (2) RIMDB costs (customer-side), (3) development of RIMDB open-source solutions, (4) regulatory and government influences, (5) network and connectivity development, and (6) increase in number of sensors. Because each of these six key drivers can develop into different directions, future projections were created for the key drivers, serving as the foundation for the three final scenarios. The scenarios combine different future projections for the key drivers in a narrative fashion and are set in the year 2025. As mentioned, the scenarios only provide a range of possible futures, but decision-makers are advised to monitor the key driver development closely over the coming years in order to be prepared in time for potential changes in one direction or another.

In the first scenario, a cloud-based, open-source RIMDB is made available to the community, decreasing the market share of the proprietary RIMDB vendors significantly. RIMDB vendors should be aware that such developments have taken place in other software domains, and prepare themselves for that case by monitoring the competition and considering new business models.

In the second scenario, the further diffusion of cloud systems, alongside strong governmental support for the opportunities that big data has to offer, foster a golden age for data. Big data technologies, including RIMDB, are now also affordable for very cost-sensitive businesses such as SMEs. This strengthens the overall economy as the value from data can be reaped more effectively. Many industrial players rely on SMEs. To ensure that the latter do not lag behind when it comes to adapting to big data analysis due to their limited resources, decision-makers concerned with Industry 4.0 should evaluate these developments closely and enable SMEs to take part in the Industry 4.0.

In the third scenario, a new platform business model for RIMDB is described: One vendor makes the code of its RIMDB freely available to the community, and uses it as a platform on which developers can write solutions for use cases relating to big data. Platform users can purchase these solutions. The vendor makes money as soon as a developer sells his/her solution through a share of the revenue. Considering that eventually, market saturation for real-time big data analysis systems will be reached if they become the state of the art, this business model may be particularly interesting for RIMDB vendors to consider for sustained competitive advantage.

List of Figures

Figure 1: Generic Value Chain; highlighted: Technology and Operations	9
Figure 2: Google Trends for “Internet of Things” (blue) and “IoT” (red)	14
Figure 3: Google Trends for “Big Data (yellow), “Internet of Things” (blue) and “IoT” (red).....	15
Figure 4: Value Delivery in the IoT (own illustration, based on Fleisch et al., 2014, p.7, Fig. 4)	16
Figure 5: Technology Life Cycle Stage	34
Figure 6: Google Trends for “SAP HANA” (blue), “TimesTen” (red), “Hekaton” (yellow) and “IBM DB2” (green)	37
Figure 7: Key drivers and projections (visualisation in EIDOS)	46
Figure 8: Consistency matrix (visualisation in EIDOS)	50
Figure 9: Scenario cluster selection	51

List of Tables

Table 1: Overview over Expert Sample and Interviews	22
Table 2: Overview over Interview Rounds.....	24
Table 3: Summary of Strengths and Weaknesses of the Scenario Method.....	30
Table 4: The 31 Drivers; highlighted in bold are the key drivers, described in section 4.1.5.2	43
Table 5: Key Drivers and Projections (Descriptions)	49

List of Abbreviations

BTLE	Bluetooth Low-Energy
CF	Corporate Foresight
CPS	Cyber-Physical Systems
CRM	Customer Relationship Management
DB	Database
e.g.	exempli gratia
EC	European Commission
ERP	Enterprise Resource Planning
et al.	et alii
EU	European Union
i.e.	id est
IM	In-Memory
IoS	Internet Of Services
IoT	Internet Of Things
IT	Information Technology
LiFi	Light Based High-Speed Communication Technology
n.d.	No Date
N/A	Not Available
NoSQL	Not Only Structured Query Language
OF	Operational Foresight
OLAP	Online Analytical Processing
OLTP	Online Transactional Processing
p.	Page
R&D	Research And Development
RFID	Radio-Frequency Identification

RIMDB	Relational In-Memory Database
SCD	Smart Connected Device
SF	Strategic Foresight
SME	Small/Medium Enterprise
SQL	Structured Query Language
TER	Technology Exploration Report
TF	Technology Foresight
TM	Technology Management
US	United States (Of America)

1 Motivation, research question and thesis outline

1.1 Problem statement: Real-time decisions to cope with change in the Industry 4.0

“The pace of change will never again be as slow as it is now” (Fjord, 2015, p.3), stated Matthew Bishop (an editor for The Economist) at the 2015 Innovation Forum. For companies, change means uncertainty, which can pose a serious threat for their survival (Zahra, 1996; Porter, 1985b). Hence, the acceleration of change means that the uncertainty in the environment of companies increases as well. Porter (1985b) notes that technological change can have a particularly significant influence on a firm’s environment. Moreover, technology is of particular significance to strategy-making because it can influence the entire value chain (Porter, 1985a). Therefore, according to various authors, the main task of strategic management is to anticipate what will happen in the future in order to make decisions now that can prepare their company for future change, thereby reducing uncertainty (Drucker, 1986; Grant, 2010; Winterfeld & Karnani, 1987). Considering the increasing pace of change, strategic management will also have to find a way to make decisions faster. In order to respond to change fast, real-time strategic decision-making for competitive advantage will increasingly become relevant for managers and other decision-makers (Rohrbeck & Bade, 2012).

One technological trend that can have a significant influence on strategic decision-making at the moment, and also a well-known buzzword, is the Internet of Things (IoT): For example, Google Trends show a rather steep incline in search volumes for the IoT since around 2013 (see Figure 2, p.14), and Gartner, Inc. already rated the IoT as the most hyped emerging technology in 2014 (Gartner, Inc., 2014a). The IoT can be defined as connectivity of *anything* or *anyone* at any *time* or *place* with *anything* or *anyone*, and in an ideal case with any *path* or *network* and any *service* (Perera, Zaslavsky, Christen, Georgakopoulos, 2014, p.416). These *things* that characterise the IoT are also called smart connected devices (SCD) (Porter & Heppelmann, 2014). Such an SCD can be any object that has been fitted with a sensor or actor, making it electronic and capable of collecting and sending data (Atzori, Iera & Morabito, 2010; Porter & Heppelmann, 2014; Swan, 2012). Highlighting the enormous potential impact of the IoT, Gartner, Inc. (2014b) forecast an increase in SCD from 4.9 billion in 2015 to around 25 billion until 2020. Other authors’ estimations go even further; e.g. Swan (2012) estimates 50 billion SCD by 2020. The very large amounts of data (also called big data) that the SCD collect can then be analysed for various purposes (Jankowski, Covello, Bellini, Ritchie, Costa, 2014, p.4-7; Porter & Heppelmann, 2014, p.68). Since merely accumulating data cannot provide much insight, data analysis is one central aspect where value can be created in the IoT (Fleisch, Weinberger & Wortmann, 2014). Today, most of the potential that lies in big data is not exploited fully as it is simply collected and stored, but not analysed further (Manyika, Chui, Bisson & Woetzel et al., 2015). Regarding the significance of the IoT and big data for companies, the area of Industry 4.0 (also called the Industrial Internet of Things) is given particular attention by researchers, practitioners and policy-makers (see

e.g. Hermann, Pentek & Otto, 2015; Koch, Kuge, Geissenbauer & Schrauf, 2014; Manyika et al., 2015), as elaborated in more detail in section 2.2.3 (pp.16). Firstly, this has to do with estimations that the value from the IoT is expected to be the largest by billions of US-Dollars in production environments, such as factories (Manyika et al., 2015). Thus, the economic impact of IoT and big data on industrial settings will likely be extensive and can influence both companies individually and the economy as a whole. Secondly, since industrial settings are concerned with the production of goods and services, the value chain element operations (Porter, 1985a) plays a particularly large role. Business operations are those value-creating activities where any kind of resource or input is transformed into an output in the form of a product or service, including production and resource planning (CSCMP & Sanders, 2014; Porter, 1985a; Robbins & Coulter, 2012). Operations are a crucial part of all day-to-day management tasks relevant for product and service delivery, and can have a significant impact on the bottom-line performance and financials of a company (CSCMP & Sanders, 2014). Due to this high relevance of operations management, many businesses aim to improve the latter by implementing novel technologies (Robbins & Coulter, 2012). That significant value could be added to the management of operations by incorporating real-time data from the IoT is thus not too much of a stretch. Considering the high relevance of operations for industrial settings, the particular focus on Industry 4.0 therefore indeed makes sense. In order to conduct such real-time analyses of large amounts of data, novel technologies are required. A technology for this purpose are relational in-memory databases (RIMDB). Instead of storing data on a separate hard-disk, and moving it to the main memory for analysis, in an RIMDB data can be both stored and analysed directly in the main memory of the database (DB). Thus, the use of in-memory (IM) technology can tremendously speed up data analysis processes, and real-time (or near real-time) data analysis is made possible (see Grandpierre, Buss & Esser, 2013; Krueger, Grund, Tinnefeld, Eckart, Zeier & Plattner, 2010; Loos, Lechtenbörger, Vossen, Zeier & Krüger et al., 2011; Vom Brocke, Debortoli, Müller & Reuter, 2014). RIMDB are reliable, meeting ACID criteria, (Edlich, n.d.) and best suited for storing and analysing structured data (such as tables, tags, indexes, reports) using SQL (Kanimozhi & Venkatesan, 2015). Alternative and complementary technologies in the context of big data analysis exist (see section 4.1, pp.31), e.g. NoSQL DB for unstructured data (such as audio or video files; Kanimozhi & Venkatesan, 2015). There are various examples for RIMDB, commercially distributed by large software vendors, including Oracle TimesTen (Oracle Corporation, n.d; Oracle Corporation, 2015), SAP HANA (SAP America, Inc., n.d., a), IBM DB2 with BLU (IBM Corporation, n.d.) and Microsoft Hekaton (Diaconu, Freedman, Ismert, Larson, Mittal et al., 2013).

Because of their ability to analyse large amounts of data in real-time, RIMDB can be a highly relevant enabler for real-time decision-making. A thorough understanding of the technology can help decision-makers in the Industry 4.0 to deliberate whether or not to invest in the technology in the first place, where to exploit synergies with other technologies and how it can be implemented. That is why a particular focus is being put on the technology in this thesis. Considering the disruptive changes that

the IoT can bring about, RIMDB are also subject to change and influencing forces that could shape the development of the technology. The future development of RIMDB, a big data analysis technology, could hence also have a significant impact on the development of Industry 4.0. The aforementioned other big data analysis technologies are, however, not disregarded but also considered as alternative and/or complementary to RIMDB throughout this thesis. The scope of a master thesis would not allow for an equally in-depth consideration of all big data analysis tools.

1.2 Thesis goal and research question: RIMDB future developments and Industry 4.0

Understanding the future development of RIMDB, and what these mean for Industry 4.0 as a highly relevant area in the context of big data analysis, can thus improve the decision-making of practitioners from different backgrounds. This thesis aims at guiding strategic decision-makers –faced by the increasing pace of change and expected disruptive developments in connection with the IoT– by describing possible alternative scenarios for the future development of RIMDB. Due to the emphasised heightened relevance of the IoT for industrial settings, i.e. Industry 4.0, this thesis moreover aims to provide insights of particular pertinence to strategic decision-makers in industrial settings. Additionally, a target group of high importance for this thesis is the RIMDB vendors who distribute their systems, such as Oracle, SAP and Microsoft. As the vendors of the technology, they can play a significant role in influencing the future development of RIMDB with their own strategic decisions. In order to achieve the above, this thesis will answer the following research question: *How can relational in-memory database (RIMDB) developments by 2025 affect Industry 4.0?*

The research question consists of two sub-questions:

- (1) *What are future developments of RIMDB by 2025?*
- (2) *What do these future developments mean for Industry 4.0?*

The year 2025 was set as the time frame for the thesis based on Coates' (2010) definition that foresight studies shall go more than five years into the future, but a time frame of more than ten years likely yield results of only questionable validity (given the research design and means).

This thesis is conducted in cooperation with the Technical University of Berlin, the University of Twente, and the University of Potsdam.

1.3 Laying the foundations: Relevant literature and the research design

As the “missing link” (Strebel et al., 2007, p.26, Abb.4) between engineering science and management, technology management (TM) deals with enabling technological potentials in organisations in order to achieve strategic and operative goals (Strebel et al., 2007). To achieve sufficient understanding of TM in the context of strategic decision-making, this thesis firstly reviews literature on competitive strategy, including Porter's value chain (1985a, 1985b), and the particular role of technology for competitive strategy. As already mentioned, Porter (1985a, 1985b) highlights the relevance of technology –and particularly information technology (IT)– as a cause for uncertainty and in strategic contexts.

As this thesis aims to determine future developments of a technology, namely RIMDB, a useful instrument for this purpose in the context of TM is technology foresight (TF) (Strebel et al., 2007; Schuh & Klappert, 2011). TF encompasses the identification, observation, assessment and interpretation of potentially influencing factors for change in the development of a technology, and thus serves to prepare for making more informed decisions (see Rohrbeck, Arnold & Heuer, 2007; Rohrbeck, Battistella & Huizingh, 2015). Literature on TF thus also provides a relevant foundation for this thesis, enhancing understanding of the context and chosen research design. The relevance of Industry 4.0, relating to the value chain element business operations and the disruptive technological trend IoT, was already explained above. Therefore, the literature review is rounded off by delineating the meaning of these terms and all the aspects involved in creating value in the IoT.

Due to the newness of the topic and the fact that neither a scenario analysis nor another type of foresight study on the future development of RIMDB has been conducted to the best knowledge of the researcher, the research design for this thesis is of explorative nature. As a technology foresight (TF) study, this thesis applies the scenario method, based on the steps described by Reibnitz (1988), but also drawing from other relevant literature (including Godet & Roubelat, 1996; Mietzner & Reger, 2005; Schoemaker, 1995; Van Notten, 2006). Reibnitz' approach was slightly adapted to the context of this thesis. Scenarios do not serve to predict an exact future, but to open the space of possibilities that are considered for strategic decisions by providing several, alternative, possible scenarios. To support the scenario building process, thirteen semi-structured, qualitative interviews were conducted with ten experts in four different rounds between November 2015 and March 2016. The experts were recruited through purposive non-probability sampling (Saunders et al., 2009) using the researcher's and thesis supervisors' personal and professional networks as well as online research. The experts come from various backgrounds in research and practice in connection with RIMDB and big data analysis (see section 3.1.2). The entire scenario building process was aided by the software Parmenides EIDOS™ (EIDOS).

1.4 Theoretical & practical contribution

Several aspects regarding the relevance of this thesis were already mentioned throughout the previous sections. For the reader, it is important to understand that in the continuum between basic and applied research (as illustrated by Saunders, Lews & Thornhill, 2009, p.9) this thesis tends more towards the direction of applied research. Hence, the purpose of this thesis lies mainly within the area of improving understanding of a management problem. The findings will be particularly relevant to managers in organizations. Specifically, this thesis aims at guiding decision-making of managers when it comes to decisions surrounding real-time analysis of large data volumes, as already elaborated. Because the IoT is prognosticated to have a strong impact on the competitive environment across industries, "(...) a new set of strategic choices (...)" (Porter & Heppelmann, 2014, p.66) will emerge. As described, companies will need to move towards real-time management by implementing new

technologies that can enable real-time decision-making through big data analysis. Since RIMDB are a crucial technology for enabling the latter, decision-makers need to thoroughly comprehend the technology. This thesis shall serve to provide a sufficient overview over RIMDB to enhance understanding of the technology. As no technology stands by itself without external influences, RIMDB are also subject to the changes that IoT-related developments, the technology's competitive environment and other external drivers can bring about. Thus, the future development of the technology could be shaped by various factors surrounding the developments in the IoT. On the other hand, as a key big data analysis technology, also RIMDB could have an influencing effect on these developments. Understanding the future development of RIMDB, as well as what these mean for Industry 4.0 as a highly relevant area in the context of big data analysis, can thus improve the decision-making of practitioners from different backgrounds. As described, this thesis is particularly directed towards decision-makers (a) in industrial settings or for whom Industry 4.0 will be highly relevant and (b) the vendors of RIMDB solutions. On top of that, this thesis can be useful for (c) the vendors of alternative or complementary technologies regarding big data analysis, and (d) all organizations considering the implementation of a big data solution.

Regarding the academic relevance, this thesis will enrich the growing body of research on IoT and big data with a structured TF study focusing on the highly relevant big data analysis technology RIMDB. Literature on TM, TF, competitive strategy, the value chain, business operations, the IoT and Industry 4.0 is reviewed mainly using a structured approach, unless a more narrative approach was required due to the topic. In consequence, the connections between these different fields are derived in a structured and comprehensive fashion, which can also serve as a foundation for future research. Moreover, this thesis adds to literature on TF by applying the scenario method in a very current context, namely big data analysis. Not only was the scenario approach by Reibnitz implemented, but also adapted to the context. The latter may provide guidance for future research in a similar direction. Furthermore, from the results of this thesis, future directions for research emerge.

1.5 Thesis outline

The next chapter of this thesis elaborates on relevant literature regarding technology management and competitive strategy, the role of IT for the value chain, technology foresight as a technology management tool and the value chain element business operations in the context of the IoT (Industry 4.0). The relevance of the literature for the research question is summarised at the end of the chapter. The third chapter of this thesis concerns the research design and methodology. Thereafter, the results of the thesis are presented, encompassing a thorough description of the technology RIMDB under consideration of related technologies, relevant driving forces for the future development of RIMDB and the three final scenarios. The thesis closes with conclusions for RIMDB and Industry 4.0 to answer the research question in chapter five, and a discussion of limitations and future directions in chapter six.

2 Literature review

2.1. Technology management

Technology Management (TM) has been defined as the “missing link” (Strebel et al., 2007, p.26, Abb.4) between engineering science and management, i.e. managing technological knowledge within an organization. The focus is less on specific innovation processes, but on (more generic) technological developments. The relevance of TM for organizations can be derived from its function as enabler of the technological potentials in an organization, which in turn are necessary for reaching its strategic and operative goals. Therefore, the tasks of TM can be summarised as development, allocation, implementation, retention and exploitation of technologies, while ensuring a link between engineering science and management. From a strategic perspective, TM includes activities surrounding TF, planning, development, exploitation, protection and assessment. (Strebel et al., 2007; Schuh & Klappert, 2011)

Conducting a TF study using the scenario analysis method, this thesis combines technological knowledge on RIMDB, in addition to analysing the technology’s competitive environment and other strategically important aspects. Thereby, this thesis aims at fulfilling the claim of TM to link engineering science with (strategic) management.

To further embed this thesis into the context of TM, the following sub-chapters will provide a brief overview over relevant literature and define key terms for competitive strategy and TF.

2.1.1 Introduction to competitive strategy

In 1986, Peter Drucker wrote: “(...) the one thing that is certain about the future is its uncertainty, its risks.” (p.54). Yet in order to ensure the long-term survival of a firm managers have to consider both present and future: Management may make decisions for action in the present, but inevitably these actions will impact the firm’s future. Hence, Drucker (1986) concludes: “Management has no choice but to anticipate the future, to attempt to mould it, and to balance short-range and long-range goals.” (p.89). Grant (2010) underlines: “A company’s ability to adapt to changes in its environment depends on its capacity to anticipate such changes.” (p.287). Accordingly, managers need to create and implement a strategy in the present in order to prepare their firm for the future.

Porter’s concept of competitive strategy has become one of the most dominant paradigms in strategy research (see e.g. a meta-analysis by Campbell-Hunt, 2000). He considers analysing the competitive environment of a firm and its strategic positioning as central factors for creating sustained competitive advantage. Another perspective in strategy research is the resource-based view (RBV) of the firm, arguing that sustained competitive advantage depends on a firm’s resources and capabilities (e.g. Barney, 1995; Wernerfelt, 1984). Studies (see e.g. Spanos & Lioukas, 2001) have found that both industry (i.e. competitive environment) as well as a firm’s assets (i.e. resources) can contribute to firm success significantly.

However, this thesis is mainly based on Porter's perspective on competitive strategy, because in the context of a TF study, environmental factors are deemed more relevant than one firm's particular assets. For instance, this thesis will explore the various players in the field, and the potentially influencing factors (or "drivers") of the studied technology (RIMDB). This way it will be able to provide not only an overview over both the technology itself and the competitive environment today, but also over potential future directions via the TF study. Wherever relevant, resource-related aspects will nonetheless be considered in order to ensure a holistic perspective, and not be excluded out of principle. The specifics of TF are explained in detail in section 2.1.4 (pp.10).

Firm-internal factors play a role for the value chain concept by Porter (1985a; 1985b; Porter & Miller, 1985), which is used as the basic structural model of this thesis, and elaborated below. However, analysing the resources of a particular firm is not part of this thesis, neither is the ideal composition of resources to gain sustained competitive advantage under given circumstances.

The following sections therefore proceed to explore relevant literature on competitive strategy by defining key terms, presenting prominent concepts and relating these to TM and TF.

2.1.2 Competitive strategy and the firm's environment

Competitive environments seem to be becoming less predictable and less stable (Grant, 2010), and therefore more uncertain. Zahra (1996) defines uncertainty as the "(...) dynamism, heterogeneity, and hostility of the company's environment." (p.192). Grant (2010) states: "When the firm is buffeted by unforeseen threats and where new opportunities are constantly appearing, then strategy becomes a vital tool to navigate the firm through stormy seas." (p.18). Winterfeld and Karnani (1987) note that the main concern of strategy is the future, which is why „(...) the strategic context of a firm is always uncertain (...)“ (p.187), with varying degrees and sources of uncertainty depending on the firm and its context. Competitive uncertainty and externalities which relate to the competitive environment are one key aspect of uncertainty according to the authors. Porter (1985b) particularly stresses increasingly rapid technological change as one main driver of competition in a firm's environment: "Of all the things that can change the rules of competition, technological change is among the most prominent." (p.60).

For Porter (2008), "Competitive strategy is about being different." (p.43). He defines competitive strategy as the combination of different activities with the goal of delivering "a unique mix of value" (p.43). This can both mean that a firm's activities differ from those of rivals, or that the same or similar activities are executed in a different way than those of rivals (Porter, 1996; Porter, 2008). Accordingly, strategic positioning is defined by Porter (1996) as the attempt to "achieve sustainable competitive advantage by preserving what is distinctive about a company" (p.60).

Porter (1980) therefore considers the firm's environment and more specifically its industry (or industries) central to strategy formulation: Competitive strategy is effective if it can generate a "defendable position" (p.29) towards the competitive forces in its industry/industries by taking either

“offensive or defensive action” (p.29). For Grant (2010), competitive strategy answers to the question “How should we compete?” (p.19, Fig. 1.4). Similar to Porter, he emphasizes that strategy is the link between the firm and its competitive environment.

Summarizing definitions of competitive strategy, Grant (2010) notes that (1) achieving certain goals, (2) allocating resources according to strategically relevant actions, and (3) ensuring “*consistency, integration, or cohesiveness* of decisions and actions” (p.16) are part of common strategy definitions. This summary somewhat integrates aspects of the RBV, because it also includes the allocation of resources depending on what is deemed the right strategic action in a specific competitive situation.

However, sustainable competitive advantage does not always result from a unique competitive position, because it may be imitated by other firms (Porter, 2008). Porter (1985a) describes competitive advantage in general as the gains from providing value for buyers to be higher than the cost of creating this value. Grant (2010) even goes further by stating that the purpose of business itself is the creation of value, firstly for customers, and secondly by extracting customer value in a way which allows profiting from it. Firms can achieve superior value than competitors through lower prices for equivalent benefits (cost leadership strategy), or unique benefits for which buyers are willing to pay more (differentiation strategy) (Porter, 2008; Grant, 2010).

Contrasting Porter’s view that strategies are deliberately planned and executed is that of Mintzberg (1978; 1987; 1994), who argues that strategies are always on a continuum between deliberate and emergent: An intended strategy may be planned deliberately at a certain point in time, but very rarely is it executed to the end as realized strategy without an emergent deviation from the original plan due to changes in the environment and necessary adaptation of the strategy (Mintzberg, 1978). In other words, Mintzberg’s (1978) understanding of strategy is not that of a fixed plan with systematic execution, but more that of an emergent process with a more or less clearly defined intended strategy at the outset. This notion comes from the observation that exact forecasting of the future in any competitive environment, especially on the long-term, is impossible. The more uncertain and prone to unexpected changes an environment is, the more an emergent strategy may simply be necessary for the firm to survive.

As will be explained in section 2.1.4 (pp.10), this thesis does not aim to forecast an exact future, but open the scope of possibilities that are considered when strategy is formulated. Whether this strategy turns out to be planned and executed exactly as planned, or emerges and is adapted to change in the environment will depend upon the particular circumstances and choices of any decision-makers who consider this thesis for their strategy-making. Before giving details on TF, the following section highlights the particular relevance of IT for value creation in Porter’s framework.

2.1.3 The role of IT in competitive strategy and the value chain

As mentioned, competitive strategy results from a combination of activities for value creation. To understand competitive advantage and the activities relevant to creating competitive strategy, Porter

(1985a) proposes the value chain as a tool, which "(...) disaggregates a firm into its strategically relevant activities in order to understand the behaviour of costs and the existing and potential sources of differentiation." (Porter, 1985a, p.33).

The primary activities of the value chain move from inbound logistics over to operations, from there to outbound logistics, then to marketing and sales and finally to service (see Figure 1 below). Support activities, including technology, back these primary activities, thus permeating every point of the value chain. (Porter, 1985a)

Porter (1985b) moreover sees the value chain as a means for analysing the role of technology in competitive advantage of a firm as support activity across all value chain elements. The peculiarity of information technology (IT) is that it "is permeating the value chain at every point" (Porter & Miller, 1985, p.151). The authors consider IT of crucial strategic significance.

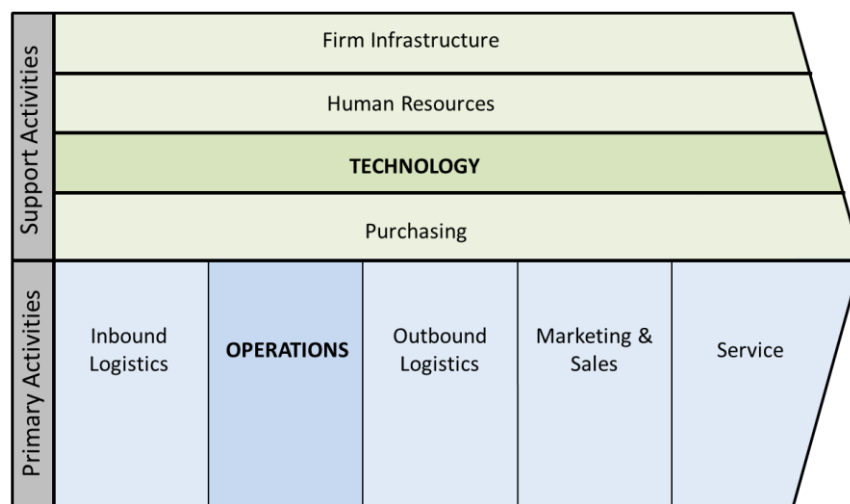


Figure 1: Generic Value Chain; highlighted: Technology and Operations (own illustration, based on Porter, 1985a, p.37)

Carr (2003) opposes this notion with his article "IT doesn't matter", and proclaims that IT has become a commodity due to its ubiquitous presence. However, he admits that technologies, especially infrastructural technologies whose value increases when shared by different players, can create competitive advantages for first-movers: "In addition to enabling new, more efficient operating methods, infrastructural technologies often lead to broader market changes. Here, too, a company that sees what's coming can gain a step on myopic rivals." (p.43). Still, Carr (2003) warns that commoditization occurs particularly fast for IT. He adds that the type of investment and use of IT are crucial factors for businesses regarding their cost and risk management. Thus, he does not negate the strategic significance of IT entirely, and especially not before the technology in question becomes a commodity. Considering its technology life cycle stage (see section 4.1.2, pp.33), RIMDB (the technology in focus in this thesis) are not classified as a commodity. Therefore, even when taking Carr's perspective, RIMDB remain relevant to strategic decision-making.

Going back to the value chain, it is important to note that one firm's value chain does not stand alone, but is embedded in a larger value system, which may incorporate the value chains of suppliers, buyers and of other channels on the way to the end customer (Porter, 1985a; Porter & Miller, 1985). Porter

(1985a) reasons: "Gaining and sustaining competitive advantage depends on understanding not only a firm's value chain but how the firm fits in the overall value system." (p.34). Since IT can exert influence and is part of all points in the value chain, the relevance of IT for competitive strategy is thus further highlighted, especially considering that this realisation of its importance will also hold true for all points in the larger value system in the competitive field. In addition, Porter (1985b) sees technology as one main driver of changes in industry structure and industry boundaries.

Other scholars have expanded the value chain model with the value shop and value network as ways of configuring value creation in different contexts (see e.g. Stabell & Fjeldstad, 1998). However, irrespective of the exact configuration of elements, for the purpose of this thesis the notion that IT permeates all value creation elements shall be accepted. Moreover, the objective of this thesis is not to conduct a value chain or value system analysis, but the value chain, and more specifically its value creating elements, shall serve as basic structural models for this thesis and are important for understanding the relevance of the research question: If IT permeates all primary value creation activities and is particularly relevant to strategic positioning, then IT developments and innovations can exert an influence on all value creation activities (and also the larger value system). Thus, analysing such IT developments in detail shall be a high priority for strategic decision-makers. This thesis aims at providing insight for the latter to current developments in the field of real-time big data analysis through the technology RIMDB, as further elaborated below.

Considering the relevance of technology for value creation and the necessity to anticipate environmental changes, a specific technology strategy for firms suggests itself. Technology strategy can be defined as "(...) the company's approach to deploying its technological resources to seize opportunities in the market and to counteract uncertainty." (Zahra, 1996, p.192). Porter (1985b) also considered the creation of a technology strategy very relevant for strategic decision-makers, and therefore provides seven generic steps for the formulation of technology strategy.

As mentioned, accurate prediction or forecasting of the future are nearly impossible, but thinking about what might happen in the future is necessary when creating strategy (Grant, 2010). This is where TF, as a means for formulating technology strategy, comes into play: TF can better prepare firms for what may happen in the uncertain future and thereby reduce the competitive uncertainty to a certain extent. The following section will go into more detail on TF. One specific tool that can be used for a TF study, and is also used in this thesis, is the scenario analysis, which is described further below in the methodology chapter (pp.19).

2.1.4 From technology strategy to technology foresight

To understand in how far TF can aid a technology strategy, it is firstly important to define what scholars understand under the concept of foresight. Miles (2010) describes "(...) 'Foresight' as a popular way of describing broad programs of study of research and innovation plans and priorities in the light of potential long-term future developments." (p.1449). Martin (2010) specifies that TF,

concerning foresight activities on technological developments, has become a widely used term in different areas¹. Rohrbeck, Battistella and Huizingh (2015) note that the research stream corporate and strategic foresight has grown remarkably over the last decade. In their literature review of over 30 years of research on future orientation of organizations, Rohrbeck and Bade (2012) found that over time, previously distinct research streams seem to have been integrated into the terms corporate or strategic foresight as the most common terminology for future research in an organizational context. Rohrbeck and Schwarz (2013) summarize that the terms corporate foresight (CF), strategic foresight (SF) and organizational foresight (OF) can be used as synonyms, which will also apply to this thesis. Although this thesis is a TF and not a SF study, explained further below, it is still relevant to provide an overview over SF to understand the context and purpose of this study.

Rohrbeck et al. (2007, p.4) combine market aspects (competitive intelligence, consumer foresight, political environment foresight) with the technology-centric perspective by defining SF as

“(...) the identification, assessment and usage of weak signals² to recognize and give warning about threats and opportunities at an early stage. Sources of weak signals are the political, socio-cultural and competitive environments as well as science and technology. Strategic Foresight defines the methods, the actors, the process and the system needed to enhance the competitive position of a company.”

In consequence, not only information on the technology itself but also its environment is of relevance to SF activities. The latter should then serve to strengthen the competitive position of a company. In a later article, Rohrbeck et al. (2015, p.2) also emphasise this point, i.e. the creation of competitive advantage through foresight activities, and define CF as follows:

„Corporate foresight permits an organization to lay the foundation for future competitive advantage. Corporate Foresight is identifying, observing and interpreting factors that induce change, determining possible organization-specific implications, and triggering appropriate organizational responses. Corporate foresight involves multiple stakeholders and creates value through providing access to critical resources ahead of competition, preparing the organization for change, and permitting the organization to steer proactively towards a desired future.”

While accepting the notion that changing environments and uncertainty are a fact for companies (especially when facing the future) the question remains why this changing environment poses such a particular threat for firms. Rohrbeck and Gemünden (2011) summarize three reasons: (a) the rate of change has increased significantly over the last years; (b) firms have been found to be ignorant towards changes occurring in their environment; (c) firms are subject to inertia, e.g. due to the complexity of their internal and external structures and prevailing cognitions. They also state how CF attempts to mitigate these challenges: (1) as initiator, CF can directly stimulate the beginning of innovation process in organizations; (2) as opponent, CF challenges prevalent ideas and assumptions the innovators may have in a company; (3) as strategist, CF can e.g. provide guidance for the creation

¹ For more details on the development of the term “Technology Foresight”, see Martin (2010) and Miles (2010).

² Weak signals can serve to prevent strategic surprises, i.e. “(...) sudden, urgent, unfamiliar changes in the firm's perspective which threaten either a major profit reversal or loss of a major opportunity.” (Ansoff, 1976, p.22). Weak signals can announce such discontinuities. For more background on weak signals see Ansoff (1976).

of strategy and exploration of possible new business models. Although not embedded in a corporate context, a TF study can still fulfil these roles to an extent if considered in the strategic decision-making of a firm.

Rohrbeck and Bade (2012) relate CF to Igor Ansoff's idea that "(...) strategic management will move towards real-time and competitive advantage will be built on the basis of timely responses to external (discontinuous) change (...)." (p.2). Especially considering the possibilities that new technologies (including RIMDB and other IoT-related technologies) offer for real-time data analysis and the relevance of technology for sustained competitive advantage in combination with increasingly rapid technological change, the necessity for moving towards real-time strategic management does not appear far-fetched.

Investigating SF empirically, Rohrbeck and Schwarz (2013) could find rather scarce support for value contribution of SF in the majority of the examined firms (in part due to methodological difficulties), although all previously determined value contributions were found to be achievable. Enhanced perception (either through novel insights into environmental changes or reduced uncertainty) turned out to be the most significant value contribution. In a previous study, Rohrbeck (2012) found empirical evidence for several areas where CF was indeed able to create value. For example, companies were able to identify relevant change, support breaking away from path dependency and trigger innovation initiatives through CF activities.

These studies (and other more recent case studies, see e.g. the special issue on CF of *Technological Forecasting & Social Change*, Vol. 101, 2015) already provide first evidence that CF does add value to companies, especially with regard to the formulation of competitive strategy. Future research will have to further investigate in this direction to broaden the empirical evidence base.

Other research has examined in how far SF can enhance the agility and dynamic capabilities of a firm, and has found positive effects (see e.g. Vecchiato, 2015). This shall only be mentioned here for completeness sake, as no particular firm is the focus of this thesis and dynamic capabilities are therefore not a central topic.

A slightly different perspective on foresight, and more specifically SF, has been provided by Coates (2010), who sees foresight as "(...) an end state; it is not a process for getting to that end state, and it is not a particular instrument or tool for getting there." (p.1428). He states: "The study of the future is not and cannot be a science. It is an art form that draws on all the arts, sciences, and humanities." (p.1430). Moreover, Coates advocates that futures work need not imply any planning or management actions and restricts foresight to all activities going five or more years into the future. He names a number of popular methods and techniques for future studies, amongst others literature review, scenarios, interviews, brainstorming and key technologies. In accordance with his definition of SF as a future state, he argues that such multitude of methods and techniques only seem appropriate. Coates

(2010) also sees a variety of uses for SF, such as coping with the unexpected, widening alternatives, and stimulating creative thinking.

Some of the arguments put forward by Coates, including that foresight deals with more long-term futures (five years and beyond from the present), and that a number of methods and techniques may serve for the purpose of SF, coincide with that of other authors. These notions shall also be considered valid for this thesis, for example by setting the time frame of the scenario analysis at over five years in the future (in 2025). However, in his rejection of planning and management actions as implications from future work he differs distinctly from other understandings of foresight. For the purpose of this thesis, the definitions of SF and CF by Rohrbeck et al. (2007; 2015) shall be applied, as they exhibit a stronger link with the notions of management literature and strategic planning.

Because this thesis is not part of the strategy formulation of one specific company, but aims to guide decision-making for strategy in corporate as well as other contexts (such as policy-making), this thesis is not a *corporate* foresight, but a *technology* foresight study. Unlike a CF study, this thesis will not go as far as formulating strategy, but aims to provide valuable insights and recommendations which can provide the *basis* for strategy formulation of decision-makers. Apart from that, the principles for a TF study are the same as those elaborated above for CF.

To do so, this thesis applies the scenario method (see section 3.2, pp.24), a foresight instrument, for the technology RIMDB. From the resulting scenarios, conclusions are drawn on how potential future developments of RIMDB may affect business operations as part of the value chain.

Above, the relevance of the value chain for competitive strategy and the relevance of IT for all value chain points were already established. The following sections will now go deeper into the aspect of business operations, one key element of the value chain, and elaborate on its particular relevance to this research.

2.2 Business operations, the Internet of Things and Industry 4.0

2.2.1 Overview over business operations

Porter (1985a) defines the value chain element *operations* as “Activities associated with transforming inputs into the final product form, such as machining, packaging, assembly, equipment maintenance, testing, printing, and facility operations.” (p.40). Robbins and Coulter (2012) describe *operations management* as “(...) the transformation process that converts resources into finished goods and services” (p.516). The authors indicate that resources can be diverse inputs from technology, equipment and materials over people to information and capital – similar to inputs from Porter’s (1985a) support activities in the value chain (firm infrastructure, human resource management, technology development and procurement).

Because of this transformative role from inputs to goods or services, the operations of a business can have a direct impact on many decisions, activities and ultimately costs which is why operations are the central function many businesses “(...) go to in order to improve performance and the financial bottom

line” (CSCMP & Sanders, 2014, p.3). To reach high levels of efficiency and productivity, which in turn influences costs, one important aspect of operations management are the identification, planning and coordination of resources – from a technological perspective usually solved through Enterprise Resource Planning (ERP) software systems (CSCMP & Sanders, 2014).

Although all elements of the value chain are vital for a firm, focus in this research is put on the element operations, firstly because of the central role of operations concerning day-to-day management tasks relevant for product/service delivery, including resource planning and production. It shall also be mentioned here that the technology in focus of the scenario analysis, RIMDB, are considered notably utile in the context of operations (explained in depth further below; see section 2.2.3, pp.16). Secondly, within the means of this master thesis, not all aspects of the value chain can be explored in equal depth considering the time- and resource-intensity of a scenario analysis and given constraints. Thus, it was deemed more important to explore one key element of the value chain in depth than merely scratch at the surface of all value chain elements.

Similar to Porter’s (1985) notion of technology as a value chain support activity, Robbins and Coulter (2012) highlight that, considering current market pressures for timely delivery of goods and services to customers, “Smart companies are looking at ways to harness technology to improve operations management.” (p.526). In particular, the authors mention controlling of costs through “(...) predictive maintenance, remote diagnostics and utility cost savings.” (p.526). The latter keywords already point towards sections 2.2.2 and 2.2.3, which deals in more depth with the link between operations and current key technological trends (Internet of Things, Big Data and Industry 4.0).

2.2.2 The Internet of Things and value creation

One of the most discussed technological developments impacting businesses across industries is the Internet of Things (IoT): Gartner, Inc. already rated the IoT as most hyped emerging technology in 2014 (2014a), and Google Trends show a continuous incline in search volumes for the IoT since around 2013 (see Figure 2 below).

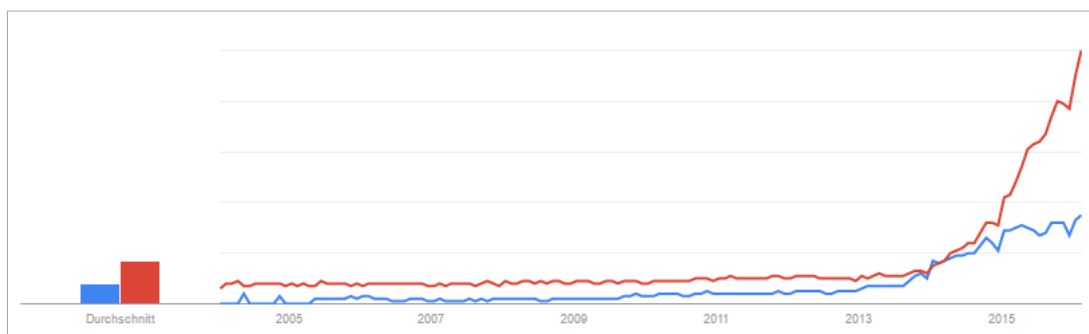


Figure 2: Google Trends for “Internet of Things” (blue) and “IoT” (red); Screenshot taken on 23.02.2016

The IoT can be defined as „(...) the pervasive presence around us of a variety of things or objects – such as Radio-Frequency Identification (RFID) tags, sensors, actuators, mobile phones, etc. – which, through unique addressing schemes, are able to interact with each other and cooperate with their neighbors to reach common goals.” (Atzori et al., 2010, p.2787; based on Giusto, Atzori, Iera &

Morabit, 2010). The authors add that the IoT could enable a vision of “(...) ‘anytime, anywhere, anymedia, anything’ communications.” (p.2803). Similarly, Perera et al. (2014, p.416) adopt a definition of IoT based on the European Research Cluster on the Internet of Things (IERC)³: Connectivity of *anything* or *anyone* at any *time* or *place* with *anything* or *anyone*, and in an ideal case with any *path* or *network* and any *service*. Although other slightly different definitions exist (see e.g. Atzori et al., 2010 and Perera et al., 2014), the latter definition shall be used for this thesis.

Kopetz (2011) points out that the differentiating factor of the IoT to existing devices with connection to the internet is the enormous prognosticated size of the IoT. According to Gartner, Inc. (2014b), the number of smart connected devices (SCD) will increase from 4.9 billion in 2015 to around 25 billion until 2020. Swan (2012) even estimates that 50 billion SCD will exist by 2020.

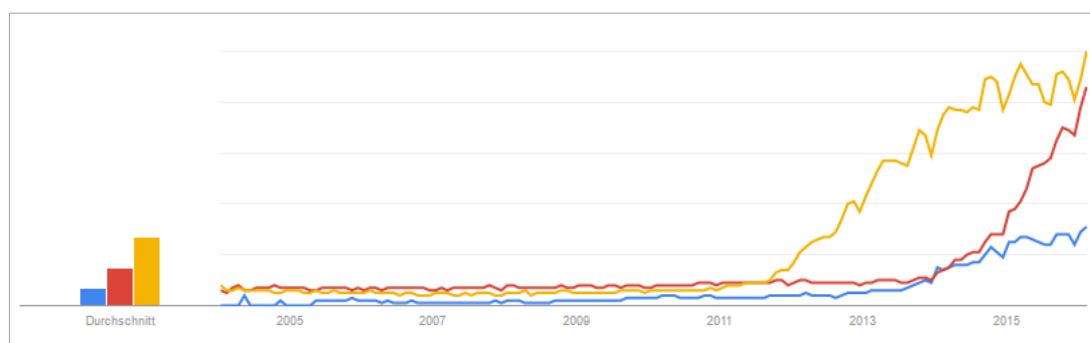


Figure 3: Google Trends for “Big Data (yellow), “Internet of Things” (blue) and “IoT” (red); Screenshot taken on 23.02.2016

On the consumer-side, SCD encompass a new generation of products, or “everyday objects that have not previously seemed electronic at all” (Swan, 2012, p. 218). On the business-side, fields expected to be widely affected by the IoT include “automation and industrial manufacturing, logistics, business/process management, intelligent transportation of people and goods” (Atzori et al., 2010, p.2787). According to Porter and Heppelmann (2014), SCD will transform the inherent characteristics of products due to integration of new IoT-related technologies, as well as how value is created. Considering the expected growth in SCD, it can be expected that the data volumes these SCD create grow extensively, resulting in a related trend: Big Data (see Figure 3 above for Google Trends). Many publications treating the topic of Big Data conclude that amassing large data volumes alone will not help business, but the *analysis* of these immense data volumes can evoke actionable insights, and at last add value (see e.g. Buhl, Röglinger, Moser & Heidemann, 2013; Lycett, 2013; McAfee & Brynjolfsson, 2012; LaValle, Lesser, Shockley, Hopkins & Kruschwitz, 2011).

Data analysis is also described as one of the five layers of value creation in the IoT (see Figure 4, p.16), identified in a Bosch White Paper based on an analysis of various IoT applications (Fleisch et al., 2014): Direct customer value is created through a physical thing⁴ (layer 1), and digital services (layer 5). The physical thing is related to the digital service via a sensor or actor (layer 2) via connectivity (layer 3) via

³ For more information on the IERC, please see <http://www.internet-of-things-research.eu/>. For more details on the IERC’s definition of the IoT, please see http://www.internet-of-things-research.eu/about_iot.htm.

⁴ Physical things in Fleisch et al.’s (2014) description correspond to the aforementioned SCD.

analytics (layer 4), which also contribute to the creation of customer value indirectly. As mentioned, this fourth layer (analysis) is of particularly high relevance for the IoT, because this is where “(...) sensor data is collected, stored, checked for plausibility, and classified” (p.7). As a result, SCD, equipped with (embedded) sensors or actors, need network connectivity in order to deliver data to an analytics system. Data analytics can then deliver results to digital services.

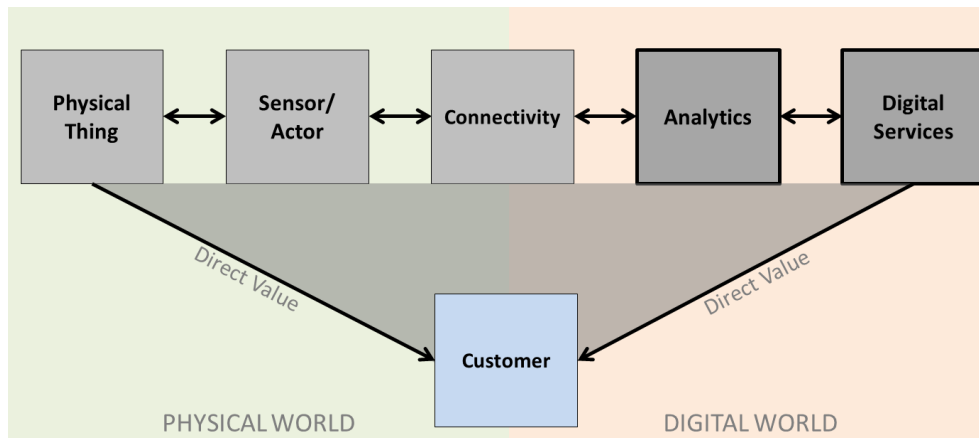


Figure 4: Value Delivery in the IoT (own illustration, based on Fleisch et al., 2014, p.7, Fig. 4)

Considering the fast-growing number of SCD in the IoT, and with it extensively expanding data volumes available to decision-makers, the increasing importance of moving towards real-time strategic management, as mentioned, shall be highlighted once more: Analytics of such large amounts of data, and in real-time, should be able to provide much deeper and profound insights to strategic decision-making than previously possible. Consequently, the digital services using insights from such Big Data analytics should provide even better value to customers. Of the data available in the IoT today, less than one percent is actually used and exploited fully (Manyika et al., 2015), further accentuating the still untapped potentials of Big Data analysis. The already mentioned RIMDB are one central technology for real-time analysis of large data amounts. The technology and potential competitors are explained further in section 4.1 (pp.31).

Developments in the IoT and concerning real-time Big Data analytics are expected to have a large impact on value creating activities. One area on which specific focus is being put by players today, are developments in industrial production and manufacturing processes: the vision of Industry 4.0. The following section gives a brief overview over this term and how it connects the IoT and business operations, and section 2.3 (p.18) concludes with the formulation of the Research Question.

2.2.3 Industry 4.0 or the Industrial Internet of Things

Especially in Germany, various initiatives from government and industry exist to support the development of Industry 4.0⁵ (Hermann et al., 2015).

⁵ See for example: <http://industrie40.vdma.org/>, <http://www.plattform-i40.de>, <http://industrie40.gtai.de>, <http://www.hightech-strategie.de/de/Industrie-4-0-59.php>, <https://www.bmw.de/DE/Themen/Industrie/industrie-4-0.html>, <http://www.siemens.de/industrie-4.0/pages/home.aspx>

Based on the definition by German Platform Industry 4.0, consultants Strategy& and PwC (Koch et al., 2014) define Industry 4.0 as follows:

“The term Industry 4.0 stands for the fourth industrial revolution. Best understood as a new level of organisation and control over the entire value chain of the life cycle of products, it is geared towards increasingly individualised customer requirements. (...) The basis for the fourth industrial revolution is the availability of all relevant information in real time by connecting all instances involved in the value chain. The ability to derive the optimal value-added flow at any time from the data is also vital. The connection of people, things and systems creates dynamic, self-organising, real-time optimised valueadded connections within and across companies. These can be optimised according to different criteria such as costs, availability and consumption of resources.” (p.10)

Koch et al. (2014) further highlight three aspects in their definition of Industry 4.0: (1) Vertical and horizontal value chains are digitized and increasingly integrated; (2) product and service offerings are increasingly digitized (“embedded systems”, p.11); and (3) digital business models are introduced. Thus, Industry 4.0 describes the vision of IoT for the industrial sector, and “The potential uses of the industrial internet go far beyond the optimisation of production technologies” (p.7). The authors consider data analytics a core capability in the Industry 4.0, but observe that although data is already being collected in some areas, the use of these data is still vastly unstructured. They also state that “Established business models will become more data-based, efficient and customer-focused due to the increased connectivity and use of data/ analytical abilities in all processes.” (p.33).

Closely related terms used to describe these developments – in particular outside of Germany – include the Industrial Internet, Advanced Manufacturing, Integrated Industry, Smart Industry, Smart Manufacturing and Smart Factory (Hermann et al., 2015). However, some of these terms are not exactly superimposable with Industry 4.0 and could also be seen as components of Industry 4.0. A Smart Factory, for example, would not go as far as encompassing the entire value chain, as per Koch et al.’s definition would be the case for Industry 4.0.

A literature review by Hermann et al. (2015) identified four key components of Industry 4.0: (1) Cyber-Physical Systems (CPS), which describes integration of processes from digital and physical world, including smart products with integrated sensors and actuators; (2) the Internet of Things, which the authors define as the network in which CPS cooperate; (3) the Internet of Services, which describes the offer and combination of services through the internet and dynamic value chain activity distribution; and (4) Smart Factories, which is characterised by context-awareness systems assisting task execution of both humans and machines. Based on their literature review, the authors moreover emphasise the high value of real-time capability (collection and analysis of data), as well as real-time communication and cooperation of CPS through the IoT for Industry 4.0, amongst other aspects.

Overall, Manyika et al. (2015) found that the expected value coming from B2B applications in the IoT will be twice as large as that from B2C applications, but add that linking B2B and B2C applications may even further enhance the expected value. The authors postulate that the impact of the IoT will be by far the largest in the setting of production environments with standardisation where repetitive work is

conducted (as such, also hospitals and farms), further heightening the relevance of Industry 4.0 from both the perspective of individual firms as well as the economy in general.

In industrial settings operations, the transformation of inputs to goods or services standardised processes, including manufacturing, production and resource planning, are one crucial area of activity. It would go far beyond the time- and resource-related constraints of this thesis to provide a complete overview on all the works recently published surrounding the topics of IoT in industrial settings and Industry 4.0. Yet, based on the above it should be clear that researchers, practitioners and policy-makers give particular significance to Industry 4.0 (i.e. the IoT in industrial settings). Accordingly, real-time analysis of large data amounts as a key mechanism for exploiting and delivering value in the IoT is also highly relevant in this setting, as mentioned. For the purpose of this thesis, industrial settings are thus considered a highly relevant area for the impact of the IoT and Big Data analysis in business operations.

2.3 Formulation of the research question

Recapitulating, TF was established as a utile means for aiding strategic decision-making. Moreover, the time frame for TF was set for at least five years into the future. Considering the means and method chosen for this thesis, a time frame longer than ten years could yield results of questionable validity, which is why the time frame for the scenario analysis in this thesis was set for the year 2025. Naturally, it would be interesting to consider all value chain elements equally, since IT permeates the value chain at all points. However, this cannot be achieved within the means of a master thesis, so the choice for one highly relevant field had to be made. Business operations, as a key value chain element, was determined as particularly relevant for analysing how the IoT and big data analysis can affect strategic decision-making in the future. Industry 4.0, as defined above, was moreover highlighted as an area where the effects of the IoT on operations can be observed particularly well. Both the strategic necessity of real-time strategic management, especially considering the increasing pace of change and uncertainty, and insights from real-time analysis of big data as a key value creating activity in the IoT were already discussed. RIMDB are considered one crucial technology which can be applied for such real-time big data analytics, and will be explained in more detail in section 4.1 (pp.31).

Considering all points above, this thesis aims at answering the following research question: *How will RIMDB developments by 2025 affect Industry 4.0?*

This research question consists of two sub-questions, which will be addressed step-by-step:

1. *What are future developments of RIMDB by 2025?*
2. *What do these future developments mean for Industry 4.0?*

The next section will expand on methodological considerations of this thesis. Later sections will give an overview over the technology RIMDB, and continue with the scenario analysis as well as conclusions for Industry 4.0 by the year 2025.

3 Research design and methodology

Because the topic is very current and no scenario analysis or other type of foresight study on the central big data analysis technology RIMDB has yet been conducted to the best knowledge of the researcher, the methodology of this thesis is of explorative nature. This section firstly explains the procedure of data collection, including a description of the literature review as well as expert sample and interviews. It also gives an overview over the scenario method in general before describing the approach to scenario analysis in this thesis including its strengths and weaknesses.

3.1 Data collection procedure

3.1.1 Combination approach of systematic and narrative literature review

For the purpose of this master thesis, relevant literature was reviewed for the sections building on previous knowledge. Given time and resource constraints of a master thesis, the literature review could not be exhaustive. For all sections, Web of Science was used as a primary source for online literature research. The academic search engine Google Scholar provided a secondary source. Also, the Business Source Complete database and its search function proved to be a useful tool. Further, recommendations by the supervisors of this thesis and literature from the course “Technology Management” by Prof. Dr. Hans Georg Gemünden at TU Berlin and by Prof. Dr. Katharina Hölzle University of Potsdam (summer term of 2014) served as guidance. Generally, the literature review followed the approach of selecting publications of higher quality and higher relevance to the thesis instead of a very large quantity.

Technology Management and Technology Foresight

The following keywords were used for online research: technology management, technological management, technology foresight, technological foresight, corporate foresight, strategic foresight. Because of a strong theoretical and conceptual focus in this section, publications were excluded from the search results, based on the following criteria: (a) pure case studies; (b) publications focusing on one foresight method not used in this study (e.g. Roadmapping, Delphi); (c) publications from sources of questionable impact and/or reputation; (d) publications published before the year 2000 with less than 100 citations; (e) publications published between 2000-2008 with less than 50 citations. For publications newer than 2008, less relevance was given to citations, and more relevance to the reputability of the source and author(s). The general criteria for publication selection were as follows: (I) if several publications were very similar in the content relevant to this thesis, those publications with more citations were preferred over those with less citations; (II) publications from journals with higher impact factor were preferred to those from journals with lower impact factor if the content of the publication relevant to the thesis was very much similar; (III) if a publication built on and extended previous literature, more recent were preferred over less recent publications (unless the less recent publication included definitions or frameworks of particular relevance); (IV) review articles were preferred over separate articles (unless the separate articles were needed for completeness or a more

in-depth view). Moreover, focus was put on René Rohrbeck, a reputable scholar in the field who has published a number of articles with high citations. Most articles considered in this thesis relating to the topics TF and TM originate from the journal *Technological Forecasting and Societal Change*. Search in very high impact management journals, such as *Academy of Management Review* or *Academy of Management Journal*, did not yield useful results for this topic.

Competitive Strategy and Competitive Advantage

Literature relating to competitive strategy and competitive advantage, both introductory and in the context of TM, was searched using the keywords strategy, strategic management, competitive strategy, competitive advantage, value chain, creating value, value creation, technology strategy, technological strategy, technological management, technology management, as well as combinations thereof. Publications were excluded from the results if (a) the scope was too narrow (e.g. if only effects of IT on dimensions of performance were analysed), (b) the focus of the publication did not fit the topic of the thesis (e.g. pure focus on technology policy), (c) the article had less than 100 citations. Publications were selected preferably if (I) they stemmed from a high impact management journal; (II) had over 300 citations (unless particularly fitting to the context based on scanning of the abstract – then publications with 100-300 citations were accepted). Because this strand of literature includes many basic articles, no particular time frame was set for selection of publications. A strong focus was put on publications (including both books and articles) by Michael E. Porter, as his framework was selected as the basic structural model of this thesis.

Business Operations, the Internet of Things and Industry 4.0

For the sections on business operations, the IoT and Industry 4.0, the keywords value chain, value creation, business operations, Internet of Things, IoT, Big Data, Industry 4.0, Industrie 4.0, smart production, smart manufacturing, smart factory and smart factories were searched by themselves and in various combinations. Moreover, literature already found for the previous sections was used as well, especially those describing business operations more generally. Criteria for selection of publications in these sections were less strict and the review followed a more narrative approach, because value creation in the IoT, especially with a focus on business operations, are relatively new and explorative topics. A systematic literature review would have been too restrictive for researching this explorative context; the narrative literature review enables the exploration of a wide range of avenues and possible sources of information, and adapt the research when new information is revealed (see Bryman & Bell, 2007). To enrich the search outcome, non-academic publications (e.g. popular journal articles, white papers, reports from consulting agencies, other online references) were also searched for via the search engine Google, and in part selected as literature for these sections.

Scenario Method

Methodology-focused literature on scenario analysis was searched using the keywords scenario analysis, scenario method, scenario approach and scenario planning. Publications were mainly

excluded from the results (a) if they consisted of a case study or mere application of the methodology; (b) the source and/or authors were of questionable reputation; (c) if they had fewer than 100 citations. Accordingly, publications were selected: (I) with a minimum of 100 citations; (II) when similar in content, publications published in journals with a higher impact factor were preferred; (III) when similar in content, more recent publications were preferred over older articles (particularly regarding review articles; unless the less recent article was more relevant). Moreover, the search focused specifically on publications by Mietzner and Reger as established authors in the field. In addition, the references of already found publications for TM and TF served as a source of information.

Technology Exploration Report (TER)

The information for the TER on RIMDB was collected from a number of sources. For similar reasons as in the section on business operations, the IoT and Industry 4.0, the literature review for the TER was conducted in a more narrative fashion than the other sections, and did not only include academic literature from top journals, but also popular journal articles, reports and white papers, blog articles, other online references and publications from the hardware providers of RIMDB. Especially with the latter group of references, care was taken to not only refer to literature from one provider to ensure a comprehensive overview. Keywords for the explorative search included various combinations of in-memory database and main-memory database, as well as SQL⁶, NoSQL⁷, relational database, relational in-memory database, and big data analysis. In addition to the literature review, the TER relied on the first round of expert interviews, especially for assessing the technology life cycle stage of RIMDB.

3.1.2 Semi-structured expert interviews and the expert sample

This section contains information relevant to the methodological characteristics of the expert sample and interviews. Section 3.2 describes the scenario process in more depth. Aiming at supporting the scenario approach, qualitative, semi-structured expert interviews were conducted in four phases, allowing for much deeper insights than e.g. a quantitative questionnaire could provide. The approach was deemed most suitable due to the explorative nature of the thesis, because it is possible to vary the order of questions, adapt the way questions are posed due to context, or to add questions in order to inquire more deeply when necessary. Considering the complexity of the topic and context, these possibilities could ensure more valuable and deeper insights (see Saunders et al. 2009).

Sampling Process and Sample Composition

The expert sample was recruited using purposive non-probability sampling (Saunders et al., 2009): The search for suitable experts was conducted through the researcher's personal and professional networks, extensive online research⁸ and the expert network of Prof. Dr. Katharina Hölzle's chair at

⁶ SQL means Structured Query Language, the standard language for relational database management systems; See <http://www.sqlcourse.com/intro.html> for more information

⁷ NoSQL means Not Only SQL and is used to describe a group of non-relational and largely distributed database systems; See <http://www.planetcassandra.org/what-is-nosql/> for more information

⁸ E.g. via the professional social network LinkedIn and a Google search for suitable companies and departments

the University of Potsdam. The goal was to find a heterogeneous group of experts in the field of RIMDB, knowledgeable of Big Data analysis and the IoT, with either a technical or business perspective, consisting of researchers from different institutions and practitioners with different backgrounds. Overall, these goals were achieved. At the outset, efforts were undertaken to recruit experts throughout Germany – albeit rather unsuccessfully, so that most of the search was concentrated around Berlin.

ID	Expert (anonymized where required)	Interview Round	Interview Date	Interview Modus	Transcript/ Notes
E1	Jochen Schröder. Chief Solutions Architect at T-Systems; lead for various projects relating to RIMDB and other IoT and Big Data technologies	R1	27.11.2015	Telephone	Attachment A, pp.69
		R3	18.02.2016	Face-to-face	Attachment C, pp.121
E2	Researcher in the field of software architecture	R1	27.11.2015	E-Mail	Attachment A, pp.75
E3	Lars Francke. Freelance consultant around and implementation of Big Data, Hadoop & Spark projects.	R1	28.11.2015	Telephone	Attachment A, pp.79
E4	Jörg Frank. Consultant at Syncwork, an international consulting firm for software and IT solutions with an office located in Berlin; Director of a division dealing with database and other big data solutions	R2	29.01.2016	Face-to-face	Attachment B, pp.87
		R3	15.03.2016	E-Mail	Attachment C, pp.130
E5	CEO and Co-Founder of a German start-up, providing an enterprise application based on a relational in-memory database; previously obtained a PhD in the field of in-memory database systems	R2	29.01.2016	Face-to-face	Attachment B, pp.96
E6	Research Assistant and PhD Student in the field of in-memory data databases and related topics	R2	04.02.2016	Face-to-face (together with E7)	Attachment B, pp.101
		R3	22.02.2016	E-Mail	Attachment C, p.132
E7	Research Assistant and PhD Candidate in the field of in-memory database technology and data streaming	R2	04.02.2016	Face-to-face (together with E6)	Attachment B, pp.101
E8	Professor for Database Systems and related topics at a German Technical University of Applied Sciences; member of research and expert groups on big data analytics.	R2	08.02.2016	Telephone	Attachment B, pp.112
E9	Researcher at a German research institute for intelligent software technologies; part of a research group on database systems in cooperation with other institutes	R1	25.02.2015	E-Mail	Attachment A, p.83
E10	Lars Geißler, CEO and Co-Founder of WebXells GmbH	R4	23.02.2016	Face-to-face	Attachment D, pp.133

Table 1: Overview over Expert Sample and Interviews

By extensive search, in total a selection of ten experts (out of sixteen contacted) were willing to participate. Three of the experts participated in two research phases each, amounting to a total of thirteen expert consultations. Most of the contacted experts who did not wish to participate in the study named lack of time as their reason for declining. It appears that experts in the field of big data

analysis are currently very much in demand, which is logical due to the popularity and currency of the topic, and thus seem to get rather swamped with similar requests for interviews and workshops such as the one for this thesis. Moreover, experts who were not familiar with the method of scenario analysis (or in general, technology foresight) sometimes regarded the mere principle of a future study with strong scepticism. Unfortunately, despite a thorough explanation of the method as best as possible, the researcher was not able to dissolve their scepticism in all instances, further explaining why several experts declined the invitation to an interview. Table 1 above gives an overview over the sample, interviews dates and modus, notes and/or transcripts in the attachment, and who was consulted during which research phase(s).

The Expert Interviews

The interviews in this thesis were conducted using three different modi: (1) When possible, face-to-face, as it offers the richest contextual information and more room to explore topics (e.g. by showing print-outs of relevant information in a sequential manner). (2) If the expert was not available for a face-to-face interview, a telephone interview was suggested. (3) Only if neither (1) nor (2) modus was possible, e-mail was used.

Each face-to-face or telephone interview lasted between around 25 minutes and 55 minutes – depending on the research phase and both conciseness and communicativeness of the experts. As suggested by Saunders et al. (2009), open questions were asked to describe the context, and probing questions to entice more in-depth and critical replies. As the authors also suggest, audio files were recorded when possible in order to transcribe the interviews later. These were backed up by written notes in case of a recording failure. In some cases, the interviewees did not wish to be recorded, or no means of recording was at hand due to contextual circumstances (e.g. very spontaneous decision on an interview appointment). In these instances, the researcher could only rely on her written notes. Attachments A, B, C and D (p.68-146) include the interview notes and verbatim transcripts as well as interview guidelines and additional materials. Please note that the interviews were all conducted in German, as all interviewees as well as the researcher are German and this facilitated a more eased conversation. For the results section, all relevant information provided by the experts was translated into English. Please note that some of the experts were only willing to participate in the study under the promise of anonymity (see Table 1, p.22).

The Interview Stages

In the first stage (R1), the preparation of the TER, telephone and E-Mail were more practical due to unavailability of the experts for face-to-face interviews. However, in R1, these modi were sufficient as the interview guideline (see Attachment A, pp.68) for this stage was more structured than later rounds, aiming at determining the Technology Life Cycle phase of RIMDB. The assessment of E9 was provided to the researcher at a later time during the research process, but confirmed the assessment of the other experts consulted during R1, and could thus still be integrated.

In the second stage (R2), preparation of the key driver projections, face-to-face interviews were conducted due to the complexity of the topic. One interview took place via telephone, as the interviewee was not available for a face-to-face meeting. A catalogue of questions for each key driver was compiled in advance (see Attachment B, pp.85), and in the course of the interview questions were selected depending on the replies and communicativeness of the expert.

In the third stage (R3), one interview was conducted face-to-face, so that a print-out of the consistency matrix and projection descriptions could be shown and explained, and critical questions asked, thereby fostering a deeper understanding of the issue at hand. The second and third interviews took place via e-mail, and detailed instructions aimed at creating a similar situation. In the third case, an additional brief phone call supplemented the information provided via e-mail. The assessment of E4 in this phase was provided to the researcher after the scenarios had already been built, which is why it could only be integrated to a certain extent.

The fourth and last stage (R4) consisted of one in-depth face-to-face expert interview. Also in this case, the modus proved to be very useful: Print-outs of four previously selected scenarios were presented to the interviewee in a draft form (see Attachment E, pp.142) one after another. In each case, the expert was asked to describe his thoughts regarding a) what the scenario means for RIMDB and b) the potential influences on business operations.

Interview Round ID	Interview Round Title	Description of Interview Round	Number of Consulted Experts Per Round
R1	TER	Assessment of technology life cycle stage and additional questions on business potential and possible influencing factors	4
R2	Projections	Identification of future projections for the six identified key drivers	5
R3	Consistency Matrix	Verification and gathering of additional input on consistency matrix	3
R4	Scenarios	Verification of scenario selection, gathering of additional input on scenario background, and conclusions for Industry 4.0	1
Total Number of Expert Consultations			13

Table 2: Overview over Interview Rounds

Analysis of the Collected Data

The data analysis was facilitated by the interview transcripts, or -where no audio recording was available- interview notes. Overall, the approach was as follows: From the transcript/notes of all interviews in a particular research phase, the relevant statements were highlighted and extracted. Thereafter, the selected relevant statements were coded depending on which section of the interview guideline they belonged to, and later clustered and condensed. E.g. in R2, a list of bullet points extracted from the interviews for each key driver was created in order to derive projections for each key driver. This iterative approach of data extraction with coding and clustering was necessary, because relevant statements from the interviews did not always follow the original interview guideline and thus had to be summarized in a coherent account before making inferences and proceeding with

next steps. More detailed information on the data analysis of the expert interviews is part of the next section, explaining the scenario analysis approach in this thesis.

3.2 The scenario method

In this thesis, future scenarios for RIMDB as an essential step towards answering the research question are developed. This section provides an overview over key terminology and conceptualization of the scenario method, and presents the approach applied in this thesis.

3.2.1 Definition of key terms and conceptualization

Godet and Roubelat (1996) define scenario as “A description of a future situation and the course of events which allows one to move forward from the original situation to the future situation.” (p.166). Considering that developing scenarios is one tool for TF, a main purpose of developing scenarios is aiding strategic decision-making, amongst others for corporate managers and strategists. For the same purpose and as a useful TF-tool, the scenario method is applied in this thesis: Scenarios should help decision-makers to take various possible futures into consideration, in order to create more informed strategies. Specifically, the scenarios in this thesis are directed towards the strategic decision-makers and strategists of RIMDB vendors. They can also be useful for policy-makers and other practitioners dealing with real-time big data analysis. Scenarios do *not* serve to predict or forecast *one* future, but provide *several possible* futures and help to avoid tunnel vision on an issue by showing alternatives. (Schoemaker, 1995; Mietzner & Reger, 2005)

Various terms for the scenario method can be used synonymously, including scenario technique, scenario approach and scenario analysis (Mietzner & Reger, 2005). ‘Scenario building’ is the process of creating scenarios, while ‘scenario planning’ means integration of scenario building into corporate planning processes (Mietzner & Reger, 2005, p.224). This thesis will not go quite as far as scenario planning, as it is not embedded in a corporate context and can thus only make suggestions to decision- and strategy-makers.

Based on a literature review, Mietzner and Reger (2005) note that various definitions of scenarios exist. Moreover, scenarios can be built using various approaches and for different purposes (see e.g. Mietzner & Reger, 2005; Godet & Roubelat, 1996). Van Notten (2006) distinguishes between the more intuitive, deductive scenario development process mostly prevalent in the Anglo-American regions, and the more systematic and model-driven, inductive scenario development process of Continental Europe, which often uses computer software. He highlights that both intuitive and analytical approaches can be combined. Inputs for the scenario process can be of qualitative nature, most suitable “(...) for the analysis of complex situations characterized by high levels of uncertainty (...)” (Van Notten, 2006, p.13), or quantitative, e.g. by using computer models. Also in this case, the author points out that a combination-approach can increase consistency and robustness of the scenarios.

Within the scenario method, different scenario types can be distinguished. Godet and Roubelat (1996) distinguish between (1) exploratory (“starting from past and present trends and leading to a likely

future”, p.166) and (2) anticipatory or normative (“built on the basis of different visions of the future; they may be either desired or, on the contrary, feared”, p.166) scenarios. Börjeson, Höjer and Dreborg, Ekvall and Finnveden (2006) follow a similar approach, but add the category of predictive scenarios (attempting an exact prediction or forecast of what will happen).

Scenarios can either be possible, realizable or desirable. Possible scenarios include possible options, realizable scenarios consider possible options while adding the aspect of feasibility, while desirable scenarios may be possible but not necessarily realizable (Godet & Roubelat, 1996). There are various other conceptualizations of types of scenarios (see e.g. Mietzner & Reger, 2005 for an overview). Providing details on all scenario types would go beyond the scope of this thesis. The scenarios developed for the purpose of this thesis shall be explorative and possible. Neither aspects of desirability nor feasibility are of relevance in this thesis, as it shall serve to open up the scope of possibilities and thinking, and not provide an exact framework for working towards a highly desired future. Hence, this thesis builds scenarios for possible alternative futures based on past and present trends. Summarizing, the subject of the scenario analysis conducted in this thesis is describing possible alternative futures for the future development of RIMDB in order to move towards answering the research question.

Geographically, the scenarios are not restricted to any particular region. As reasoned previously, the time frame for the scenarios was set for the year 2025. This time frame also seems appropriate given Gartner’s (2014a) prognosis that the IoT will reach the last stage of the Hype Cycle – The Plateau of Productivity within five to ten years.

3.2.2 The scenario approach based on Reibnitz

Literature in the following section on the scenario approach taken in this thesis is based on the literature review article by Mietzner and Reger (2005) and the book on Scenario Techniques by Reibnitz (1988). As mentioned, various approaches for conducting a scenario analysis exist. As recommended by Van Notten (2006), the approach chosen for the purpose of this thesis is a combination of systematic, inductive, intuitive and deductive scenario building, and both qualitative and quantitative inputs. For the most part, this thesis relies on the systematic steps proposed by Reibnitz and Geschka, and adapted by others (e.g. Heinecke and Schwager) – amongst others because this approach suggests the creation of explorative scenarios, which were determined as the scenario type with the best fit as an aid for answering the research question in the explorative context of this thesis. Since the scenario method leaves room for adaptation of the technique to the circumstances, in some instances the scenario approach taken in this thesis deviates slightly from the original approach suggested by Reibnitz. Qualitative inputs and literature research are combined with quantitative software-aided modelling.

The computer software aiding the scenario analysis in this thesis is Parmenides EIDOS™ (EIDOS).

Step 1: Preparation of the Scenario Analysis: The TER

The first step in the approach by Reibnitz describes an analysis of the company, i.e. the task at hand including the company's strengths and weaknesses. This thesis is written at the Technical University of Berlin and University of Twente, in the context of a cooperation research project with the University of Potsdam on RIMDB, and is not embedded into a corporate context. That is why this step will be skipped. However, in preparation of the scenario analysis by gaining an in-depth understanding of the technology RIMDB, its technology life cycle stage, potential competing technologies, players in the field and the research status of RIMDB, the TER was prepared. As mentioned, four expert interviews (experts E1, E2, E3 and E9) were conducted in this stage. Moreover, the TER was based on a literature review, as discussed.

Step 2: Identification of Key Drivers

The second step described by Reibnitz will, however, be part of the methodological approach in this thesis, and will be the first step in the process for this thesis. It comprises the identification and systematic analysis of the influencing actors, factors and system dynamics. Mietzner and Reger (2005) specifically highlight one major advantage of the scenario method at this stage, as "minority votes can also be taken into account" (p.228): Not only popular or dominant views, opinions and actors considered relevant to the future development of the technology under scrutiny are included, but also potentially more peripheral factors and actors. Review and scanning of various trend reports served as the primary basis for the determination of the influencing drivers, alongside the results of the previously conducted TER and further online research. After assembling a first list, a team consisting of the author of this thesis and PhD-candidate Victoria Götz of the University of Potsdam applied the brainstorming method to narrow down and group the initial list of drivers as displayed in Table 4 (p.40-42). The resulting list of 31 drivers was transferred into the situation analysis tool of EIDOS. Thereafter, each driver was analysed by the team and compared to every other driver individually to determine how strongly each driver may influence each of the other drivers. The direction of the influence is displayed by an arrow in the software, and the strength of the influence was determined by using a scale from 0 to 3 (0=no influence; 1=weak influence; 2=medium influence; 3=strong influence). This effort resulted in an interconnected web of driver connections in the surrounding environment of RIMDB (see Attachment E, p.140). Next, EIDOS was used to create an active-passive-map based on the previous efforts. The active-passive map (see Attachment E, p.141) displays those drivers as most active which are strongly *influencing* other drivers, and those as most passive which are strongly *influenced* by other drivers. The visual representation in the active-passive map makes it easy to determine the most active, i.e. most *influencing* drivers. For this thesis, the six most active drivers were then determined as the key drivers.

Step 3: Describing Future Projections

In the next step by Reibnitz, alternative future developments (i.e. projections) for each of the identified key drivers are described. These projections are a very essential step in the scenario building

process: Each scenario is built based on a different combination of one projection per key driver. Hence, each scenario comprises six projections in this thesis – one for each key driver. To develop the projections for each key driver, a total of four semi-structured interviews with five experts were conducted (experts E4, E5, E6, E7, E8). The expert sample and interview-technique are described in more detail in section 3.1.2 (pp.21). Great care was taken for the projections in this thesis to not only be of dichotomous nature or presenting a positive, a negative and neutral option, but as qualitatively multi-faceted as possible in order to ensure the greatest possible range of alternatives for the scenario building process later in the process. Based on the interview transcripts, the information provided by the experts that was deemed the most useful for building projections was extracted and condensed in a separate document. Based on this condensed list of bullet points for each key driver, between four and five qualitatively different projections were built for each key driver.

Step 4: Defining Projection Consistency

For the fourth step in Reibnitz' approach, a consistency matrix is developed. The consistency matrix is created by conducting a cross-check of all future projections of all key drivers in pairs of two for their consistency with each other. This check serves to assess how internally consistent a scenario including both projections would be: If it does not appear plausible for two projections to occur in the same scenario, their consistency is rated lower on a scale from -3 (very low consistency) to +3 (very high consistency). The consistency matrix was first filled-in in EIDOS by the team consisting of the author of this thesis and Victoria Götz. In order to ensure the validity of the consistency matrix, three experts (E1, E4, E6) were consulted, as described in Section 3.1.2, and the matrix values were adapted accordingly in some instances. As suggested by the Reibnitz' approach, the fourth step also includes the calculation of the scenarios with suitable software, in this case EIDOS, based on the consistency matrix. The software only presents the 100 most consistent scenarios, which are grouped in clusters according to their likeness.

Step 5: Scenario Selection and Description

In step five, from each cluster of scenarios, the most representative scenario for each cluster is chosen utilizing a number of selection criteria: Based on previous literature, Mietzner and Reger (2005, p.233) name five criteria for scenario selection: (1) plausibility (there has to be the possibility that the scenarios can happen); (2) differentiation (scenarios should be distinctly different in their structure and not only vary on a few aspects); (3) consistency (the credibility of the scenarios has to be ensured through the combination of logics in a scenario and avoiding any potential inconsistencies); (4) decision-making utility (the selected scenarios should provide insights that contribute to decision-making within one single step); (5) challenge (conventional wisdom and old beliefs about what the future will look like should be challenged through the scenarios). These five criteria are applied in this thesis for selecting the final scenarios. The strongest focus was on consistency and differentiation, as the other criteria should automatically ensue due to the fact that the previous steps were taken very

carefully and precisely. This was done by comparing the most consistent and differentiated scenarios of each scenario cluster. A pre-selection of ten scenarios was then narrowed down to a selection of four scenarios (see Attachment E, p.142), while aiming at covering the largest possible scenario space. Due to the consistency values in the matrix, some projections did not appear in the list of the 100 most consistent scenarios. Again, a validity check was conducted with expert E10 (see Attachment D, pp.133), by presenting print-outs of a draft version for all four chosen scenarios to the expert, as already explained in section 3.1.2. Analysis of the insights the expert provided led to an iteration of the scenarios, so that finally, the originally chosen four scenarios were re-combined and condensed to three final scenarios. Iteration of the scenarios based on the expert's opinion poses an addition to Reibnitz's original approach. Since the sole point of validating scenarios with an expert opinion is that the scenarios may need iteration, this deviation was considered both necessary and sensible. After the three final scenarios were determined, they were fleshed out fully in a narrative fashion, using literature, statements by the interviewed experts and lastly the researcher's imagination. The narrative explanation aims at comprehensively explaining what the future looks like given the circumstances (i.e. the projections).

Step 6: Conclusions from the Scenarios – RIMDB and Business Operations in 2025

As described by the Reibnitz approach, the sixth step consists of a consequence analysis for opportunities and risks for each scenario. Due to the formulation of the research question, i.e. how RIMDB developments by 2025 affect Industry 4.0, these topics are in focus for the consequence analysis. During the interview with E10, these topics were addressed to support the concluding arguments. A set of recommendations for resulted from this analysis, and is discussed in the conclusions of this thesis. The next step as described per Reibnitz' approach as analysis of disruptive events and wild cards, did not apply in this thesis as none were identified. Lastly, the scenarios should be transferred into strategic planning. However, since this thesis is a research study and not part of corporate strategy planning, this step will be included in the development of possible actions and recommendations across all scenarios and for each specific scenario within step six in this process. The step of integration into strategic planning will have to be made by strategic decision-makers separate of this thesis, but the results may serve as a strong guidance in the process.

Although the methodological approach in this thesis was chosen and adapted carefully for answering the research question, there are certain limitations. The latter are discussed in detail in section 6 at the end of this thesis. However, before moving to the results, the following section provides a critical overview on the scenario method to enhance understanding of its strengths and weaknesses, and thus provide sufficient context for the results of this thesis.

3.2.3 A critical view on the scenario method

Drawing from their literature review, Mietzner and Reger (2005) find several strengths and weaknesses of the scenario method, summarized in Table 3 below:

Strengths	1. Instead of trying to forecast one specific future, scenarios provide a range of multiple possible futures, which can be more or less desirable
	2. Scenarios can open up the thinking of managers and corporation by showing new possibilities
	3. Scenarios may uncover previously unexpected technological discontinuities or disruptions, and weak signals, which can be incorporated into strategic planning
	4. In companies, scenarios can be used for communication purposes for opening strategic conversations
	5. Improvement of organizational learning and decision-making processes during scenario generation
	6. The scenario method is flexible and the process can be tailored to the task at hand
Weaknesses	1. The scenario method is very laborious and time-consuming
	2. Finding appropriate experts and participants for building the scenarios may be difficult
	3. The field under investigation has to be understood in depth before conducting a scenario analysis, which is why data and information on the field has to be collected and interpreted first
	4. While applying the scenario method, one may focus too much on “black-and-white scenarios” or the most desirable scenario due to wishful thinking

Table 3: Summary of Strengths and Weaknesses of the Scenario Method (Based on Mietzner & Reger, 2005, p.235-236)

A scenario analysis is very time-consuming and laborious. An in-depth understanding of the field under investigation has to be achieved first, and appropriate experts are not easy to come by. Also in the case of this thesis, with fewer time and resource restrictions, some steps could have been improved. For example, a larger expert sample size could have increased the heterogeneity of the expert sample and distribution of professions within the sample. Although two experts (E1 and E3) could be recruited despite not being located in Berlin, efforts to recruit experts from all of Germany remained mostly unsuccessful. Due to the limited resources of a master thesis, the search had to focus on Berlin and its surroundings (where the researcher was located). Similar to the sample size, more focus on recruitment in other areas than Berlin could have increased the expert sample heterogeneity. The latter could in turn have yielded more diverse and potentially even more insightful results. Nevertheless, a rather diverse expert sample was achieved considering the sample composition (see Table 1, p.23).

Because the researcher is not a trained IT expert or professional, the TER was prepared previous to the scenario analysis in order to gain more understanding of the field and technology under scrutiny. Due to consultation of experts along the way, the TER should provide a detailed-enough overview over RIMDB to ensure sufficient understanding of the technology and field.

Although researchers applying a qualitative approach will always be victim to their own subjectivity to certain extent, great care was taken to reduce any possible wishful-thinking while building and selecting scenarios. One advantage of the thesis at hand in comparison with other scenario analyses is that it is not written in a corporate context, and so may be less subject to wishful thinking than in a

corporate context. Moreover, the diversity of the expert sample involved in the scenario building process was intended to protect from such wishful thinking.

Additionally, it should be noted that the mentioned multitude of approaches and processes of the scenario method and consequential flexibility of the approach to be tailored to the specific task, which Mietzner and Reger (2005) see as a strength of the method, may also be a weakness: It will be difficult to ensure that the approach chosen is perfectly suitable for the task at hand, considering that so many variations exist. However, the researcher is confident that she has chosen the most suitable process for this thesis to her best knowledge. With this in mind, the following chapter presents the results of the TER and scenario analysis.

4 Results

The results section of this thesis is twofold: Before the scenario analysis results are presented in section 4.2, the following section gives an overview over the technology RIMDB in form of the Technology Exploration Report (TER) to enhance understanding of the context. Thereafter, chapter 5 (pp.57) concludes what these scenarios could mean for Industry 4.0 to answer the research question fully.

4.1 Scenario preparation: Technology Exploration Report (TER)

This part of the thesis describes the technology RIMDB and its research status. Moreover, the technology life cycle stage is assessed with the help of expert interviews. Then, the active players and competitive field are drafted shortly. The business potential of RIMDB is then evaluated before moving to the last point before the scenarios: the driving forces and trends.

4.1.1 Relational In-Memory Database Systems (RIMDB): Technology description and research status

Please note that this is not a thesis in the field of information systems. Thus, the technology description only describes those aspects of RIMDB relevant for understanding the results section of this thesis.

The Technology in Focus: RIMDB

It was already mentioned that RIMDB is a technology for the real-time analysis of large data volumes. Various authors emphasise the drastically increased performance of RIMDB, especially when it comes to retrieving and analysing large data volumes (see e.g. Grandpierre, Buss & Esser, 2013; Loos, Lechtenböcker, Vossen, Zeier & Krüger et al., 2011; Vom Brocke, Debortoli, Müller & Reuter, 2014).

As the name RIMDB (relational in-memory database) already shows, there are different components: (a) data, (b) relationality, and (c) in-memory technology. Before explaining (b) and (c) components, let us start at the beginning (a): Data can be:

- I. Structured (e.g. tables, tags, indexes, reports) in a specific data model, and therefore semantically explicit
- II. Semi-structured as a sub-form of structured data containing tags and/or other markers, but without the underlying data model structure
- III. Unstructured (e.g. sensor data such as audio files and streaming video) generated from both machines or humans

Different data has to be treated differently: (I) Structured data is usually managed using SQL in a relational database, (II) semi-structured data requires a non-relational model and is expressed e.g. using XML or Java Script Object Notation, but can be converted into structured data through various techniques, and (III) unstructured data can also be handled using non-relational databases, e.g. NoSQL (Not Only SQL) amongst some other tools. (Kanimozhi & Venkatesan, 2015)

Concluding for (b), relational databases (DB) manage structured data using SQL, and are neither suited for unstructured data, nor for semi-structured data (unless converted into a more structured form). Relational DB store and show data in relationships, i.e. in a table format. This makes it easy to search for specific data. (Techopedia, 2016)

In general, relational DB meet ACID criteria, thereby determining relational DB as reliable (Chapple, n.d.). NoSQL (non-relational) DB in turn do usually not meet ACID criteria, but can have other advantages and characteristics for certain contexts, and are usually open-source (as opposed to the many proprietary relational DB) (Edlich, n.d.).

Regarding (c) the in-memory aspect of RIMDB, it is important to know that there are two types of data processing: Online Transaction Processing (OLTP) for transactional data and Online Analytical Processing (OLAP) for analytical data. Transactional data can come from operational activities that are fed into the system, while analytical data are already manipulated, e.g. for problem solving, planning and decision-making. (Datawarehouse4u.info, n.d.)

In-memory (IM) DB can store data directly in the main memory instead of a hard-disk. That means data does not have to be moved to a different location to analyse it, but is retrieved directly from the main memory, thus making separation of OLTP and OLAP unnecessary. Contrarily, the more conventional hard-disk based DB systems require transferring of transactional data in batches from a hard-disk to the main memory for analysis, which is a main factor impeding the performance of hard-disk based DB as opposed to IM DB. With DB using IM technology this batch-wise transfer of data is not necessary anymore, resulting in lower latencies for data processing. Moreover, transactional data can be included in analyses immediately. As a result, IM DB can enable data analyses in (close to) real-time – an immense speed-up of data analysis processes: From hours or days with hard-disk DB, to minutes, seconds or even split-seconds with IM DB. (Grandpierre et al., 2013; Loos et al., 2011; Vom Brocke et al., 2014; Krueger, Grund, Tinnefeld, Eckart, Zeier & Plattner, 2010)

An IM DB can be relational (using SQL) or non-relational (e.g. NoSQL). RIMDB combine the aspects of relationality and IM technology (Hallenbeck, 2015).

Summarising, RIMDB are reliable (ACID) and are most suitable for collecting, storing and analysing large amounts of structured data in (close to) real-time by using IM technology.

Research Status

Codd already introduced a relational DB system in 1970. Loos et al. (2011) highlight that the basic principle of an IM DB was already put forward in the mid- to late-1980s (see e.g. Eich, 1988; Eich, 1989). Obviously, the use of RIMDB is not a completely new idea. However, until the price of main memory decreased to an affordable level rather recently, the idea was set aside (Loos et al., 2011).

During interview round R1 the experts were asked for their evaluation of the research status. E1 stated that he sees future research potential mainly with regard to costs for RIMDB, as well as the technology's diffusion and the usability of connected RIMDB applications. He considered all basic research on the technological components of the IM technology completed. E1 moreover named a number of application fields already known today and did not expect many novel developments in this area in the future either. Contrary to that, E9 saw much potential for future research regarding the use of RIMDB on new territory. In addition, E3 currently saws a lot of willingness to invest in research

and development of new applications which can be connected to RIMDB. Loos et al. (2011) furthermore state: "The area of in-memory database systems is currently characterized by a lot of R&D activity, with some systems such as solidDB (IBM), TimesTen (Oracle), and SAP HANA already being commercially available." (p.390). Thus, basic research appears to be completed, but there are still areas where future research could uncover new insights. The following section on the technology life cycle stage will come back to the Research and Development (R&D) intensity.

4.1.2 Technology life cycle of RIMDB

For the technology life cycle stage conceptualisation, literature by Sommerlatte and Deschamps (1985), Specht and Möhrle (2002), Gerpott (2005) as well as Burgelman, Christensen and Wheelwright (2009) serves as the foundation. As mentioned, the expert interviews conducted in R1 provide the foundation for the assessment below, and are displayed in a condensed way here: Figure 5 below shows the resulting classification for each indicator of the technology life cycle stage, based on the condensed expert assessments (red boxes). Additionally, the blue box shows the resulting overall determination of the technology life cycle stage: RIMDB are in-between the growth and maturity stage. Therefore, RIMDB have already reached a medium level of exhaustion of their competitive potential. Below, each indicator assessment is described briefly.

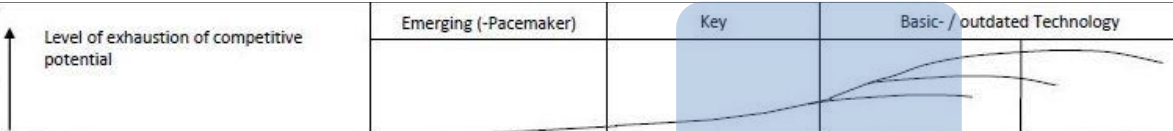
	Emerging (-Pacemaker)	Key	Basic- / outdated Technology	
Indicators	Introduction	Growth	Maturity	Aged
Uncertainty of technological performance	high	medium	low	very low
Degree of technological diffusion	very low	medium	high	very high
Number and breadth of potential application fields	unknown	increasing / big	established	decreasing
R&D intensity and development type	middle (basics)	high (applications)	low (cost reduction)	very low
Development requirements to product maturity	long term (7-15 years)	medium term (2-4 years)	short term (>1 years)	-/-
Availability of technology	very limited	waiting periods	market oriented	mass production

Figure 5: Technology Life Cycle Stage, adapted from the lecture slides of Technology Management at TU Berlin and University of Potsdam, summer 2014 (based on: Sommerlatte & Deschamps, 1985; Specht & Möhrle, 2002; Gerpott, 2005; Burgelman et al., 2009)

Uncertainty of the technological performance

There are few problems with the technology RIMDB. Some occur with the software-stacks building on RIMDB though, such as enterprise-applications, which are not yet ready to use the technology to its full extent. The speed that RIMDB can provide is therefore often nullified, e.g. when additional legacy-software is used or data is transferred into an Excel-table due to lack of better gateways to the new database. There are areas where more research is necessary to ensure smooth performance of RIMDB, e.g. its high energy consumption of main memory. Overall, performance issues can be

considered rather minor regarding the overall technological performance of RIMDB, which is why the uncertainty of technological performance can be rated as close to *low*.

Degree of technological diffusion

The diffusion of RIMDB is still lower in comparison to other (mostly hard-disk-based) DB systems. According to the experts, RIMDB are rather expensive, hampering diffusion in more cost-sensitive application areas. Many companies have to carefully assess which data is administered in-memory and which is kept on a disk that is cheaper than memory storage. However, the trend towards price declination for main memory storage appears to continue, which would support future diffusion of RIMDB. Many application fields for RIMDB appear to be untapped at the moment, offering further diffusion potential. Based on this assessment, the degree of technological diffusion as of now can be considered *low to medium*.

Number and breadth of potential application fields

Practically any field where an entire set of stored data is used as the input for a query can be a potential application field for RIMDB. They include: reporting and business analytics, data-mining and machine-learning issues such as forecasting, text-mining, evaluation of medical study data, genome analysis, geo-data-analysis, targeting and tracking of users and advertisements, recommender systems, search engines, as well as traffic planning and monitoring. Although several of these application fields are in themselves established, the main advantage of RIMDB lies in the speed with which queries can be answered. Especially when it comes to more unstructured IoT and sensor data, there are also other technologies and players aiming to grasp market share. According to the experts, RIMDB may be able to gain more ground in application fields relating to ERP and CRM (strongly relating to the aforementioned field of business operations), and other business intelligence applications, as opposed to the field of IoT data. Availability of RIMDB through cloud solutions opens application fields for more cost-sensitive players (e.g. SMEs). In sum, the number and breadth of potential application fields can be considered as *increasing/big, slightly tending towards established*.

Research and Development (R&D) intensity and development type

R&D is being conducted to further improve RIMDB performance and mitigate the performance issues mentioned above. Moreover, there are R&D activities concerning the application fields (especially when it comes to the implementation and best practices) and partly cost reduction (mostly regarding main memory storage). Overall, the intensity of R&D is *medium (neither high nor low)* as it concerns both applications *and* costs.

Development requirements to product maturity

The experts consider the RIMDB technology itself as relatively mature, as it was already implemented by several major vendors and the technological components are generally well understood. Further development requirements mainly concern further cost reduction and availability of compatible applications for the various fields where RIMDB can be used. The literature confirms this assessment,

since research has started in the 1980s and by 2011, several RIMDB were on the market (Loos et al., 2011: solidDB by IBM, TimesTen by Oracle, SAP HANA). Development requirements to product maturity are therefore *medium to short-term (less than a year to three years)*.

Availability of the technology

As a software (albeit based on hardware, unless used in conjunction with a cloud system), RIMDB are available to a large number of potential clients worldwide and not subject to the availability restrictions of more hardware-based technologies. The main impediment to availability of the technology is the cost of implementation, which is why full mass market availability cannot be inferred quite yet. This issue may be mitigated in the future due to further cost decline in main memory and increased application of cloud-technology (reducing customer costs). Another issue could be the high energy consumption of main memory in some areas. Hence, the technology is not yet fully available to a mass market, but certainly in a *market oriented* fashion (moving close to *mass market*). Aside of the understanding of the technology RIMDB itself and its life cycle stage, an overview over the active players and the competitive field of RIMDB is also important for being able to prepare the scenario analysis and determine driving forces. The following section provides this overview.

4.1.3 The active players and the competitive field

Players in the competitive field can influence the future development of RIMDB in different directions and at different points. Based on online research and the results from the expert interviews in R1, the active players relevant to the field of RIMDB can be split into six groups as described below. Please note that this is not an exhaustive list naming all players explicitly. The names of players that are cited should only serve as examples for describing the player groups and facilitate better understanding.

(1) Research institutes

Research institutes investigate the topic of RIMDB, as well as other topics related to big data analysis. For Berlin and its surrounding area alone, there are the examples of the Hasso-Plattner-Institut (HPI) (who e.g. offers an open online course⁹), the Database Systems and Information Management Group (DIMA) at Technical University of Berlin¹⁰, and the Berlin Big Data Center¹¹. Worldwide, there are many more examples, such as the Massachusetts Institute of Technology (MIT) Database Group¹² to just name one. Research in RIMDB and other big data technologies can impact the future development of RIMDB considerably, amongst others by finding solutions for the issues identified above.

(2) Governments and regulatory entities

Government organizations are relevant to the field of RIMDB when it comes to data privacy and data protection. Restrictive standards may influence RIMDB, for example regarding the use of cloud technology. For Europe, the European Commission (EC) plays a relevant role. The EC recognizes both

⁹ For more information, see <https://open.hpi.de/courses/imdb2015>

¹⁰ For more information, see <https://www.dima.tu-berlin.de/>

¹¹ For more information, see <http://www.bbdc.berlin/start>

¹² For more information, see <http://db.csail.mit.edu>

the opportunities big data could open up for the European Union (EU), as well as the potential problems with data protection that might occur (EC, 2015a; EC, 2015b). In January 2012, a reform of the rules for data protection was issued by the EC (EC, n.d.). There are also relevant national authorities, such as Die Bundesbeauftragte für den Datenschutz und die Informationsfreiheit (n.d.) in Germany, and regulatory guidelines, such as the Electronic Communications Privacy Act by the United States Congress (Doyle, 2012). Recently the international agreement Safe Harbour was declared invalid, intending to make it more difficult to send EU citizen data to the USA (Gibbs, 2015).

Additionally, governments and regulatory entities can invest into specific aspects of big data analysis, e.g. by financially supporting research and infrastructural measures, further highlighting their relevance for RIMDB.

(3) RIMDB vendors

There are a number of commercial, proprietary RIMDB products from large software vendors, for example Oracle TimesTen (Oracle Corporation, n.d; Oracle Corporation, 2015), SAP HANA (SAP America, Inc., n.d., a), IBM DB2 with BLU (IBM Corporation, n.d.) and Microsoft Hekaton (Diaconu et al., 2013). Unfortunately, no current data for market shares of RIMDB could be found. The Google Trends screenshot in Figure 6 below indicates search volumes for the named products. Because RIMDB are business tools mainly used by trained experts, and these experts may not need a Google search to research the products, the search volumes displayed by Google Trends cannot serve as a market share proxy. Nevertheless, Figure 6 can at least indicate efforts for publicity, with which SAP HANA leads the race since about 2012. The interviewed experts supported this notion.

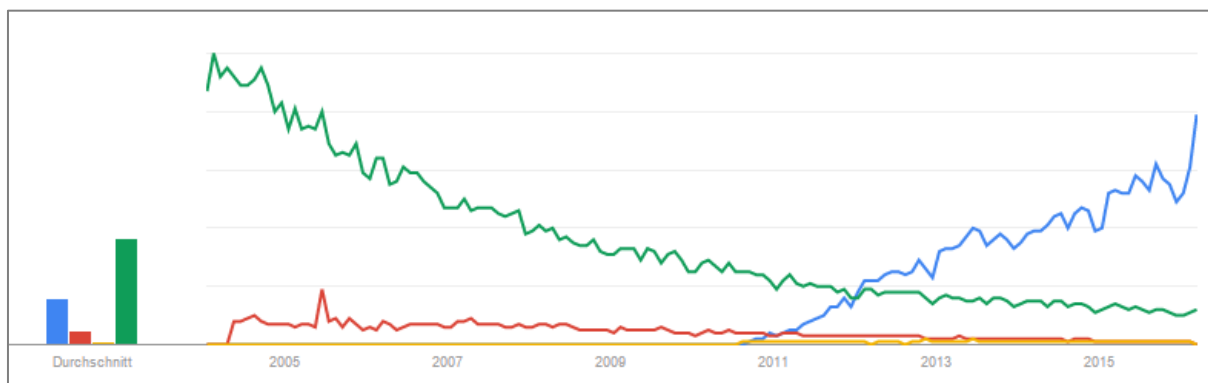


Figure 6: Google Trends for “SAP HANA” (blue), “TimesTen” (red), “Hekaton” (yellow) and “IBM DB2” (green); Screenshot taken on 08.03.2016

In addition, MonetDBLite is an open-source DB using IM technology with relational characteristics, frequently named by the experts. Other RIMDB exist, although mainly commercially distributed ones. Moreover, some DB exist that are relational, but use both IM technology and disks in combination. Examples for the latter are HP Vertica¹³ and HyperSQL¹⁴. For the purpose of this thesis, we include these DB in the group of RIMDB.

¹³ For more information, see http://www8.hp.com/us/en/software-solutions/advanced-sql-big-data-analytics/index.html?jumpid=va_udr5yqixwz

¹⁴ For more information, see <http://hsqldb.org/>

(4) Alternative and complementary technologies for big data analysis

There are a number of alternative as well as complementary technologies to RIMDB within the realm of big data analysis. As mentioned, there are non-relational/NoSQL DB, and DB that do not use IM technology, but are disk-based. That means alternative DB can (a) be using IM technology, but be non-relational (e.g. Hazelcast¹⁵, Redis¹⁶), or (b) be disk-based and either relational or non-relational. As explained, it may depend strongly on the type of data and the context which type of DB is used. Moreover, depending on both the application field in question and the strategic positioning, technologies can be either competitive or complementary to RIMDB: There is Hadoop®, a powerful open-source data processing framework managed by Apache™ which can essentially be turned into many things¹⁷. One Hadoop® project is Apache Spark™, a “(...) fast and general engine for large-scale data processing.” (The Apache Software Foundation, n.d.) which can combine various data analytics libraries. Another project is Apache Hive™ which is a data warehouse with distributed storage using its own language HiveQL (similar to SQL)¹⁸. There is also MapReduce, a system to process large datasets in parallel (Dean & Ghemawat, 2008). Many other Hadoop projects exist and are usually developed by different groups or companies (Hallenbeck, 2015). These projects can be competitors to RIMDB products for big data analysis, but Hadoop can also be used as a complementary asset to RIMDB. For example, SAP allows for use of their HANA technology together with the Hadoop framework and its solutions for additional processing power and use of e.g. NoSQL options in addition to their RIMDB (SAP America, Inc., n.d., b). Other examples for technologies that can be both competitive and complementary to RIMDBS are Amazon Web Services (AWS), a platform offering various big data storage, analytics and application solutions, and cooperate e.g. with RIMDBS vendors SAP HANA and Oracle regarding AWS cloud services (Amazon Web Services, Inc, n.d.). Also, Google offers its own cloud-based relational DB and big data analytics platform BigQuery which is compatible with other offers such as Hadoop and Spark (amongst other partners) (Google, Inc., 2015a; Google, Inc., 2015b). These alternative/competitive or complementary technologies and offers can influence the future development of RIMDB, but are also subject to similar influences as RIMDB (e.g. when it comes to regulatory restrictions for data privacy, to name just one).

(5) Potential customers of RIMDB

Larger businesses, as well as governments and any other institutes or legal entities with a need to store, process and analyse large amounts of data could be a customer for RIMDB, as described. Costs, high energy consumption and little need to process large amounts of data currently exclude many SMEs from the group of customers. This may change in the future, e.g. due to the expected cost decline. Potential customers of RIMDB are of very high relevance for the technology's future

¹⁵ For more information, see <https://hazelcast.com/>

¹⁶ For more information, see <http://redis.io>

¹⁷ Find their list of projects here: <http://hadoop.apache.org/>

¹⁸ For more information, see <http://hive.apache.org/>

development, because the willingness of customers to invest into and implement a technology (or a competing technology) could shape the market development significantly.

(6) Third-party services

There are also third party services relating to RIMDB and their distribution. Firstly, consultancies (or freelance consultants) can offer guidance when customers want to choose a DB system, and help with implementation projects. Consultancies can also provide further information and research of their own on RIMDB. Secondly, third parties can supply data either directly to the RIMDB vendor who offers the data as a package-deal to its customers, or to the RIMDB customer as an additional service. Thirdly, RIMDB rely on network and connectivity to send and receive data. For example, it is possible that a business implementing a RIMDB has branches all over the world and wants to use the technology to collect and analyse data from all these outposts in real-time – which cannot be done without sufficient connectivity. Fourthly, there are providers of applications and additional services that cooperate with RIMDB vendors. One example for the latter are the members of SAP's Startup-Focus initiative, through which SAP cooperates with start-ups offering solutions relating to big data, predictive or real-time analysis¹⁹.

4.1.4 Brief assessment of the business potential of RIMDB

With this overview over the active players and potential competitors to RIMDB for big data analysis in mind, the business potential shall be assessed in brief in this section.

Chapter 2 (pp.6) of this thesis already elaborated how IT – and thus also RIMDB – contribute to value creation across all functions in Porter's concept. Figure 4 (p.16) shows the creation of value for customers in the Io. Based on the previous sections in the TER, it should now become even clearer that RIMDB are mainly relevant for the fourth layer "Analytics" and to a lesser extent the fifth layer "Digital Services" in this illustration. Naturally, all layers are interconnected and can only function together. Grandpierre et al. (2013) highlight four general dimensions for value creation by IM technology (as compared to other DB technologies): (I) Performance (increased speed for data analysis and use of real-time data), (II) simplification (reduction of complexity and error-proneness), (III) enabling process innovations (through implementation of IM technology in business processes) and (IV) overall increased flexibility. Vom Brocke et al. (2014) confirm that the increase in information-processing capability, reduced latency and the fact that larger data volumes and more complex types of data can be analysed are main advantages of RIMDB. The bottom-line benefits on company side are that decisions can be made and trends can be recognized much faster, in addition to forecasting and intervening in working processes if need be (see e.g. Krueger et al., 2010).

In addition, there is a large potential for scalability of RIMDB: Firstly, because of the many potential application fields for RIMDB, and secondly, due to the expected growth of the IoT and prognosticated

¹⁹ For more information, see <http://startupfocus.saphana.com/>

increasing relevance of big data analysis in order to reap value from the data the IoT will create. Thus, the business potential of RIMDB can be estimated as high.

However, as also explained there are various alternative and complementary technologies and offers which could take a considerable proportion of market share. Other player groups have the power to influence the future development of RIMDB, but to a certain extent also that of these alternative and complementary technologies and offers.

Moreover, as previously stated, relatively high costs of (commercial, proprietary) RIMDB, alongside high energy consumption and lack in perceived need to process large data volumes by many companies (e.g. to low digitization), at the moment pose potential barriers to mass market implementation. This might change in the future, but can reduce the business potential of RIMDB at least in certain segments at present.

Turning to the competitive environment, it is noteworthy that most RIMDB vendors are large companies, including SAP, Oracle and IBM. Similarly, many alternative and complementary technologies are being pushed by large IT players such as Amazon and Google. This would imply that the barriers to enter the market of RIMDB (and big data analysis in general) could be rather high – especially for smaller companies or start-ups, that is, unless they collaborate with one of these big players. This should not impede the business potential for the technology RIMDB in itself, however. As explained before, various open-source solutions are also available and especially for the case of Hadoop and its projects they cannot be neglected as competitors as well.

It will depend on many different aspects whether RIMDB will manage to become state of the art for big data analysis. Within big data analysis, the question also remains which technologies of which players will mostly be used in which contexts. For example, RIMDB could mainly be used for real-time analysis of large data volumes in the realm of ERP and CRM. At this stage, all this is only speculation as the market for real-time big data analysis systems is still growing and subject to change.

Based on the above and a review of several trend reports, the following section will describe the driving forces expected to influence the development of RIMDB by 2025.

4.1.5 The driving forces for the future of RIMDB

This section firstly gives an overview over the 31 driving forces identified as influential for the future development of RIMDB. The second part of the section describes the six key drivers, and the third explains the future projections for each key driver, relevant to the subsequent scenarios.

The 31 driving forces

The drivers, i.e. the 31 driving forces, that were identified previous to determining the six key drivers from this list are shown in Table 4 (p.40-42). The 31 drivers were clustered according to relation to a specific topic into groups A-G, as shown in the table. The table also includes a brief description of each driver. The six Key Drivers and their relevance in the context of RIMDB and big data analysis are explained in more detail in the next section. Apart from the information revealed in the first round of

expert interviews and the previous sections of the TER, these 31 drivers were identified through the recent technology trend reports by Accenture (2016), EIT Digital IVZW (2016), Fjord (2015), Gartner, Inc. (2015a; 2015b), KPMG (2014) and ThoughtWorks (2015). These reports also served as the basic sources for describing the drivers in Table 4 below. In some cases, the same phenomenon was mentioned under a different name in the various reports, and a choice had to be made for naming the driver. As explained, brainstorming was used to select, narrow down and cluster the drivers.

To determine the key drivers, as mentioned, all 31 drivers were analysed for potential influences towards all other drivers, resulting in the driver grid created in EIDOS and presented in Attachment E (p.140). Due to this evaluation of the influences between drivers the researcher could create an active-passive map in EIDOS (Attachment E, p.141), thereby determining the six most influencing (active) drivers as the key drivers for this thesis. The next section describes these six key drivers in more detail.

A ENVIRONMENTAL FACTORS & ENERGY		
1	Resource Stress	Various resources are necessary to produce the increasing amounts of devices that can collect data. Overall, resource stress is prognosticated to increase over the next years and decades, and could also affect the IoT development.
2	Energy Harvesting	Most devices that are prognosticated for the IoT will need power. One very sensible option is to supply this energy through harvesting it from the surroundings of the device. This option is better suited than e.g. batteries which only have a limited lifetime.
B INFORMATION OF EVERYTHING		
3	Health Intelligence & Wearables	"Healthy is the new wealthy": Fitness trackers, wearables and similar devices in the IoT offer solutions for tracking and improving the user's health by analysing data.
4	Device Mesh	More and more different types of devices (e.g. cameras, smartphones, cars, home appliances, etc.) are connected and used for various activities (e.g. accessing application, searching for information), and increasingly cooperate. These devices collect data, thereby creating big data volumes.
5	Increase in Number of Sensors	<i>See below</i>
6	Increase in Variety of Data Sources	Because the device mesh and the IoT are increasing, the sources for data increase as well, for example video and image data, sound data and text data from devices as different as a kitchen scale, a factory robot or a streetlamp.
7	Seamless User Experience	The (digital) user experience is becoming more seamless, or "ambient" within and across devices, time and space.
8	Network and connectivity development	<i>See below</i>
C SMART ANALYTICS & SERVICES		
9	Advanced Analytics & Self-Service Delivery	Data analytics systems should enable anyone in a company to create queries and reports from data for business intelligence purposes. Analytics systems can perform advanced queries while the user interface

		and data access are simplified.
10	Advanced Machine Learning / Deep Learning	Machines are becoming more intelligent so that they can learn (without human interaction), e.g. to improve their own actions and increase their effectiveness.
11	Natural-Language Processing	Considering the multitude of devices and data sources, a core ability in the IoT can be the ability of technology to form structured information from the processing of natural language from audio or text sources.
12	Microservices & Digital Decentralization	Many different, flexible microservices serve various (changing) user needs on different IoT devices in a decentralized manner.
D	WORKFORCE	
13	Workforce Skills / Digital Dexterity	The workforce needs the skills required for applying the new technologies in the digital world so that benefits can be reaped optimally from the IoT and the data its devices create. On the other hand, some skills may not be required anymore that were important in a less digitized working context.
E	DEMAND & ADOPTION	
14	Demand of real-time response & service acceleration	<i>See below</i>
15	User Adoption B2B (Industry)	The willingness to adopt a technology by B2B users can influence various factors regarding diffusion of a technology.
16	User Adoption B2C (Consumer-Side)	The willingness to adopt a technology by B2C users can influence various factors regarding diffusion of a technology.
17	Industry Standard / Monopolization	To make the IoT work, certain standards are necessary. In many cases, these standards do not exist. On the one hand, microservices and fogging counteract monopolization in the IoT, while platformisation may create quasi-monopolies for the platform providers who can then set industry standards.
18	RIMDB costs (customer-side)	<i>See below</i>
19	Big Data Analysis Costs (General)	The overall costs for big data analysis – depending on various factors – can influence in how far users are willing to adopt a technology and drive e.g. demand, amongst others. E.g. if a business cannot afford to analyse collected data, because the necessary skills are not available and a skilled worker would be a high cost position, the business may choose to not analyse data and rely on other sources for business intelligence and improvement.
20	Regulatory and government influences	<i>See below</i>
F	COMPATIBILITIES	
21	Big Data “Big-Three” Technologies’ (Cloudera, HortonWorks, MapR) Compatibility with RIMDB	The current “Big-Three” technologies in the big data field can influence the further diffusion of RIMDB – and other technologies – by either being compatible or non-compatible with them.
22	Increased Compatibility of RIMDB with Common Big Data Technologies (Development of Hybrid Models)	RIMDB can enable compatibility to other big data technologies, thereby creating hybrid models for big data analysis.
G	BIG DATA RELATED TECHNOLOGIES	
23	Development of IoT Platforms / Platformisation	Platforms provide a certain architectural backbone

		for the IoT by aggregating devices and applications. Platform providers can set technological standards, e.g. for communication.
24	Diffusion of SQL/Relational DBs	Relational, SQL-based DB can be RIMDB, but can also be hard disk-based. This means, use and diffusion of relational, SQL-based DB systems make up one core component of RIMDB.
25	Diffusion of NoSQL DBs	NoSQL DBs are a competing technology for RIMDB, and are particularly suitable for the analysis of unstructured data. NoSQL DB are non-relational.
26	Diffusion of Hard disk-based DB Technology	Hard disk-based DB are a competing technology to RIMDB, who do not save <i>and</i> analyse data in the main memory (as RIMDB do), and still have a large market share.
27	Diffusion of Edge Computing / Fogging	Cloud computing is extended through geographic distribution of applications and services, and pre-processing of collected data, which can in turn result in lower latency (closer to real-time interaction) and increased location awareness.
28	Diffusion of Cloud Computing	Many IoT technologies are enabled through cloud computing, which makes it possible to upload and access data and information as long as network connectivity is given, store data cheaply through third-party and access applications through cloud providers.
29	Diffusion of In-Memory DB (Licensed)	In-memory technology is one key component of RIMDB, and today, RIMDB are marketed through a licensing model.
30	Development of RIMDB open-source solutions	<i>See below</i>
31	Diffusion of Hadoop Technology for Big Data Analysis	Hadoop can be both competitor and complementary technology to RIMDB.

Table 4: The 31 Drivers; highlighted in bold are the key drivers, described in section 4.1.5.2

The Six Key Drivers

As specified for the previous section, the key drivers (and all other drivers) are based on literature research and the expert interviews of R1.

1. Demand of real-time response and service acceleration

The recent Fjord trends report for 2016 (Fjord, 2015) describes how services need to accelerate in the sense that both user needs and context change dynamically and with increasing pace, which makes real-time response to these changes necessary: Smaller, faster, flatter. The authors cite Matthew Bishop (editor for The Economist, at the 2015 Innovation Forum): “The pace of change will never again be as slow as it is now” (Fjord, 2015, p.3). It was already described that real-time data analytics and -as a result from this- real-time decision-making for strategic management are expected to gain importance. However, the question remains (a) in which areas such real-time analytics and responses will actually be useful, and (b) how fast companies will recognize the need for change towards real-time response. Therefore, the actual demand of real-time response and service acceleration is determined as the first key driver for this thesis. For example, entire sectors may lag behind in adapting to increasing pace of change, or such change may be less present in specific sectors and industries. Thus, these sectors could have a lower demand for such real-time responses than other

sectors. Other aspect, such as costs for implementation of real-time response systems, or infrastructural requirements may also limit demand to a certain extent. On the other hand, if the IoT develops as expected, demand for such real-time response systems and the acceleration of services would continue to increase over the next years. As E1 pointed out, cloud solutions can (currently) not achieve actual real-time, only near real-time. For this thesis, the exact latency which is still considered real-time was not defined closely, in order to leave room for various interpretations during the interviews. Near real-time solutions, e.g. cloud systems, with low latencies are thus also meant when speaking of real-time for the purpose of this thesis.

2. RIMDB costs (customer-side)

As mentioned, according to the expert assessment, the high costs for RIMDB are currently an issue for some areas which might limit diffusion of the technology – at least in certain areas. Costs or prices for goods or services are closely related to demand as per the basic principle of demand economics (Mastrianna, 2010), further highlighting the relevance of costs of RIMDB for its future development. Costs are thus closely related to demand, and other demand-related drivers. Moreover, cost leadership was mentioned previously as one strategy companies can pursue: If one RIMDB vendor can offer a similar product for a lower price it is likely to be able to gain competitive advantage over the other vendors. There are also reasons why costs could decrease. One advantage of cloud computing can be its inexpensiveness compared to the purchase of hardware (Javadi, Anderson & Cappello, 2009). This consideration is already being applied with most RIMDB solutions. Moreover, the experts repeatedly mentioned decreasing hardware costs. Overall, RIMDB costs on the customer-side were determined as one key driver for the future development of RIMDB.

3. Development of RIMDB open-source solutions

As mentioned, most RIMDB are proprietary, and sold or licensed by commercial vendors at a relatively high price at the moment, but some open-source/ freely available RIMDB exist. Open-Source can be defined as “(...) software in which the source code is freely available for others to view, amend, and adapt. Typically it is created and maintained by a team of developers that crosses institutional and national boundaries. As such, open source software can't be appropriated by one large proprietary vendor.” (Poynder, 2001, paragraph 6). There are also other open-source big data analysis solutions, such as the Hadoop framework and its projects. The experts highlighted that further diffusion of RIMDB open-source solutions could have a highly relevant impact on the overall development of RIMDB. For example, in interview round R2, E6 explained that well-functioning open-source solutions are generally a good development for all markets, because they can set a high standard that commercial vendors need to at least fulfil, if not exceed. Moreover, there are various examples where previously proprietary software was later distributed under a free software license. Such an example is the Java platform from Sun Microsystems (Lee, 2006) which is now distributed for free since around ten years. In other cases free open-source software has developed in parallel to commercial,

proprietary solutions, such as in the case of Linux (Poynder, 2001). Based on the driver assessment, the development of RIMDB open-source solutions was determined as a third key driver for this thesis.

4. Regulatory and governmental influences

In the active players section of the TER, the role government and regulatory institutions can play was already elaborated. Such regulatory influences can touch a wide range of areas when it comes to big data analysis. As described below, the diffusion of sensors/IoT devices as well as collection, transfer and analysis of data collected by such devices can all be shaped by regulations for data privacy and security standards. It is therefore not surprising that regulatory influences became a key driver for the future development of RIMDB. The Fjord trend report emphasises how for example the EU's General Data Protection Regulation could lead to the concept of "privacy by design" for applications and digital services (Fjord, 2015). Standards regarding interfaces and intersecting points between IoT devices are also imaginable as one aspect of regulatory influences. Moreover, governments could invest in certain areas relevant to big data analysis, such as supporting education of a skilled workforce for the use of such software, or support of infrastructural measures such as network development (see below).

5. Network and connectivity development

Relating back to Fleisch et al.'s (2014) five layers of value creation in the IoT (see Figure 4, p.16), connectivity is the intersection between digital and physical world. Perera et al. (2014) mention the (semi-)automated connection between sensors and applications as an important future challenge regarding the IoT. Jankowski, Covello, Bellini, Ritchie and Costa (2014) cite ubiquitous wireless coverage as one of the main IoT enablers.

Swan (2012) highlights some technologies for data transmission, including Bluetooth Low-Energy (BTLE), WiFi and 2G/3G/4G. An article published by Techradar also presents so called Bluetooth 4.2 Beacons as highly relevant for the IoT, e.g. for sending real-time notifications (Carter, 2015). A recent Trend Report by EIT Digital IVZW (2016) considers the 5G network for the IoT as a potentially ground-breaking macro trend. The authors expect 5G to reach maturity by 2020, and offer crucial benefits for the IoT, such as higher network speeds, lower latencies and lower energy consumption. Further broadening of WiFi networks, e.g. through initiatives such as the Open Wireless Movement²⁰, as well as connection of previously network-less areas, e.g. through projects including Google Loon²¹, can play a role for network and connectivity development, as Swan (2012) mentions. Other innovative technologies may also influence the network and connection development. One example is LiFi, which "(...) refers to light based communications technology that delivers a high-speed, bidirectional networked, mobile communications in a similar manner as Wi-Fi." (pureLiFi, 2014, paragraph 1). LiFi could thus boost network speeds tremendously, and Apple is currently rumoured to start supporting LiFi on their mobile devices soon (Daws, 2016). Clearly, network and connectivity development is

²⁰ For more information, see <https://openwireless.org/>

²¹ For more information, see <https://www.google.com/loon/>

closely connected to many other driving forces related to the IoT and Big Data, and was determined as the fifth key driver for future developments of RIMDB.

6. Increase in number of sensors

Very broadly defined, a sensor collects input or information from the physical world and transforms them into output, usually in the form of an electronic signal (see e.g. Zook & Schroeder, 2005). Swan (2012) is concerned about the battery life of sensors, and found sensors to be still too expensive in 2012 for the IoT. However, Jankowski et al. (2014) consider the drop in sensor prices of almost 50 percent within the last ten years as one key enabler of the IoT. Sensors moreover appear as one layer in Fleisch et al.'s (2014) five layers of value creation in the IoT (Figure 4, p.16), as the intersection through which data moves from a physical thing (i.e. IoT device) through connectivity towards data analysis systems. Remembering Azori et al.'s (2010) definition of the IoT, and Gartner, Inc.'s (2014a) projections for the growth of SCD, of which sensors are a key component, it is clear that sensors are a central aspect necessary for the existence and growth of the IoT. In other words: Without sensors, no IoT is possible. Thus, the sixth key driver for the area of Big Data analytics, and with it RIMDB, is the increase in number of sensors. Considering its hardware components, the delta in change of hardware price for sensors may for example affect the extent of the increase in sensors. This driver is also closely connected with regulatory aspects and it is imaginable that regulatory restrictions on data privacy and security may inhibit sensor growth to a certain extent.

Future projections for the six key drivers

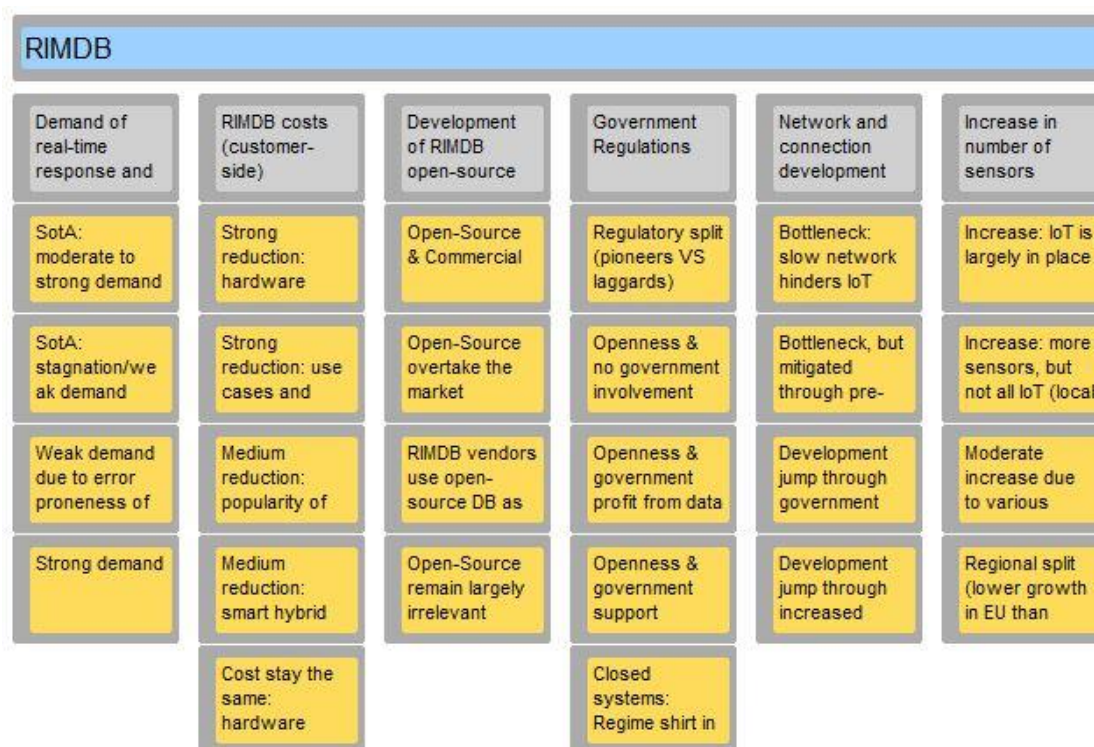


Figure 7: Key drivers and projections (visualisation in EIDOS)

Before moving to the scenarios, this section presents the future projections resulting from the second round of expert interviews (R2). Figure 7 above shows a visualisation of the key drivers and future

projections in EIDOS. The future projections were cross-checked for their consistency in the consistency matrix as shown below. It is important to understand that the consistency matrix does not display direct influences or correlations between the projections, and there is no direction of influence from one projection to the other. The question one has to pose when looking at one of the fields in the matrix is: How consistent would a scenario be that includes these two projections? The more consistent the scenario would be, the more positive the rating, and the less consistent it would, the more negative the rating. Table 5 (p.47-48) includes descriptions of all projections for the six key drivers.

KEY-DRIVER	PROJECTIONS
1. Demand of real-time response and service acceleration	1. Real-time response and service acceleration have become the state-of-the art (I): Moderate to strong demand. Due to market pressures to adoption even those firms who do not want to adopt real-time systems are forced to do so; the slow, larger corporations and laggards are still trying to catch up, but the development was not fast enough to kick all laggards out of the market – yet)
	2. Real-time response and service acceleration have become the state-of-the art (II): Stagnated/weakened demand. Those who haven't implemented the technology were kicked out of the market, and the firms who are still in the market already have implemented technologies; improvements in response speeds are only marginal and therefore do not further strengthen the demand; only start-ups/new businesses need to install new real-time systems
	3. Weak demand due to error-proneness of systems. Big data analysis systems are prone to errors, which weakened the demand for real-time systems
	4. Continuously strong demand for real-time response due to continuous improvements on response speeds and continuously growing relevance of predictive analytics and maintenance in the IoT (more "real"-time in 2025 than before)
2. RIMDB costs (customer-side)	1. Strong overall cost-reduction due to decrease in hardware costs
	2. Strong overall cost-reduction due to more use cases and thus lower project implementation costs
	3. Medium overall cost reduction due to growing popularity of cloud solutions
	4. Medium to strong overall cost reduction due to new smart hybrid systems. Hot/warm/cold data approach: Smart DB systems automatically know which data should be stored in RAM/ SSD (hot data), fast disks (warm data) or slower disks (cold data)
	5. Stagnation in costs: Hardware becomes cheaper in general, but larger data volumes require higher efficiency (increased efficiency of systems leads to stagnating prices)
3. Development of RIMDB open-source solutions	1. Open-source and commercial, proprietary RIMDB exist next to each other: Open-source solutions caught up to commercial solutions regarding both their technological abilities and market share. These well-functioning open-source RIMDB set a high standard for all commercial products.
	2. Open-source solutions overtake market. One reason for this is that a large big data company makes a well-functioning RIMDB available to the community.
	3. The large RIMDB vendors start using their RIMDB open-source systems as a vehicle for selling other systems (such as ERP- and CRM-systems)
	4. Open-source RIMDB remain largely irrelevant for the market as they cannot catch up to standard of commercial, proprietary solutions
4. Regulatory and government influences	1. "Regulatory Split" between different regions: There are some Big Data pioneers with open regulatory frameworks for exploiting data for business (e.g. USA, China) and some Big Data laggards with strict regulatory frameworks which make exploitation of data for business difficult (e.g. D)

	2. Shift towards complete openness in all regions (I): Governments largely keep out of the story and do neither inhibit nor support big data exploitation
	3. Shift towards complete openness in all regions (II): Governments of the world's developed countries have regulations which support exploitation of big data, while also using data themselves to support the government's interests (e.g. industry espionage)
	4. Shift towards complete openness in all regions (III): There are no regulations to handicap business with data, and even financial support to ensure optimal exploitation of big data, e.g. by financing research and education in the field and supporting the network and connection development
	5. Complete regime shift towards governmental control: The majority of developed countries install closed up systems; exploitation of data is supported and data is being collected, but each country/region keeps to itself, which handicaps the development of large-scale connectivity in the sense of the IoT
5. Network and connectivity development	1. Network as "Achilles' Heel" or "Bottleneck" (I): The network & connectivity development was too slow and hinders a large-scale implementation of the IoT
	2. Network as "Achilles' Heel" or "Bottleneck" (II): The network & connectivity development was too slow and hinders a large-scale implementation of the Internet of Things, but the effects of this are mitigated through pre-processing of data within IoT devices to reduce the burden on the network
	3. Network makes a jump in development due to a large-scale government investment into network & connectivity development
	4. Network develops faster than in the past due to large-scale private investments: Speeding up the network & connectivity development due to increased competition between new private investors and incumbent investors
6. Increase in number of sensors	1. Overall large increase in sensors: Large potentials in various fields: "data is the new gold"; prognoses of Gartner, Inc. etc. are true and the IoT will grow
	2. Overall large increase in sensors, but not all sensors are real IoT-solutions: The IoT is not as far developed as prognosticated; not all sensors are part of the IoT, some are merely local solutions to problems
	3. Overall marginal to medium increase in sensors: There are large potentials, but the growth in sensors was inhibited by various factors (e.g. lack of standardisation in gateways; network as "Achilles' Heel"/bottleneck; strict government regulations)
	4. Different development in different areas: Lower sensor growth in Europe (specifically: Germany) due to stricter government regulations based on the prevailing ethical values than in other areas (USA/Asia); overall net increase in sensors, but the development depends on the region

Table 5: Key Drivers and Projections (Descriptions)

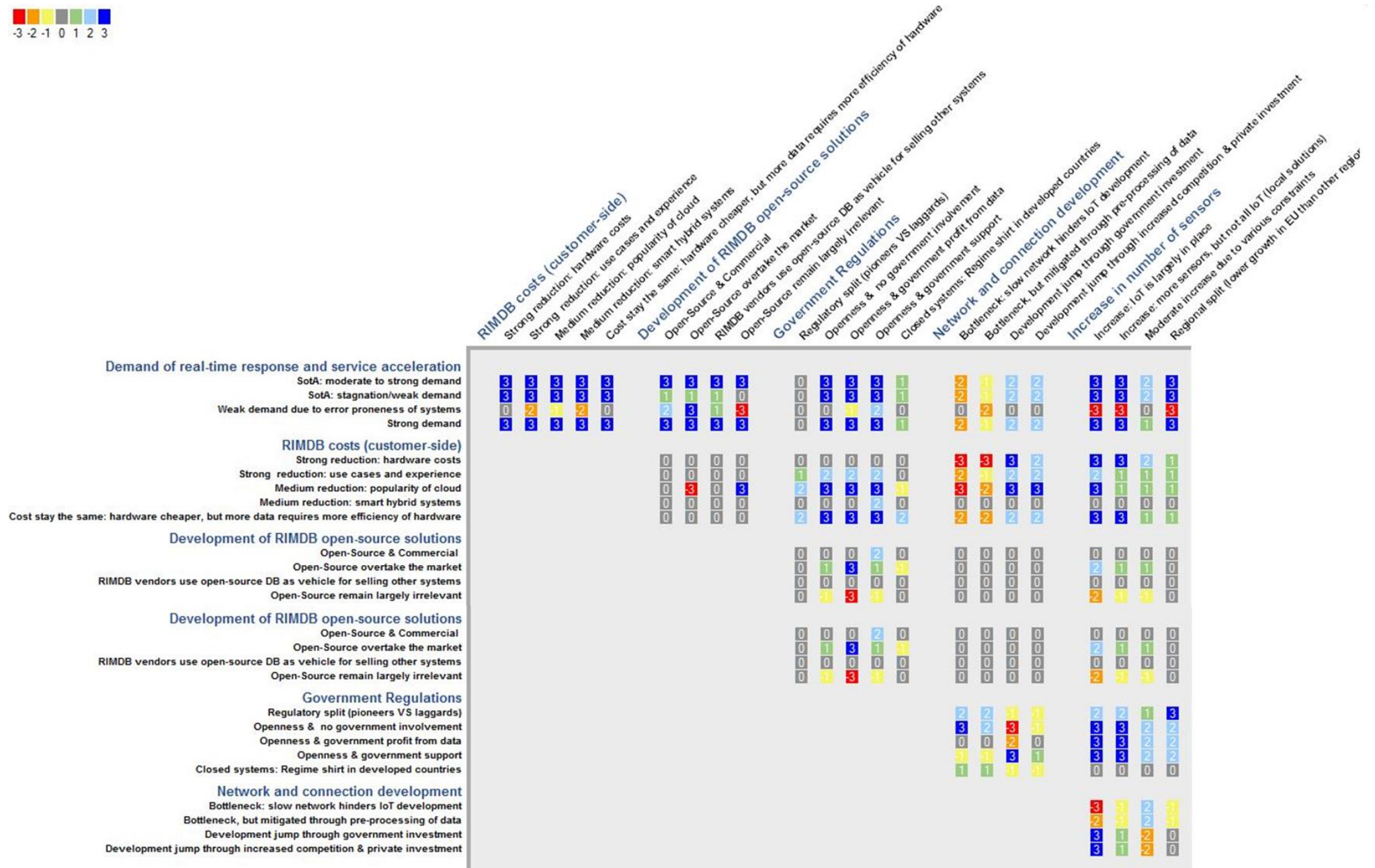


Figure 8: Consistency matrix (visualisation in EIDOS)

4.2 Scenarios

The consistency matrix (Figure 8, p.49) provides the basis for the scenario calculation in EIDOS. The software presents only the 100 most consistent scenarios. Similar scenarios are clustered more closely together, dissimilar scenarios are further apart, as can be seen in Figure 9 below. Interestingly, several projections did not appear in any of the 100 most consistent scenarios listed in EIDOS. This can be attributed to the ratings in the consistency matrix. The projection “weak demand due to error proneness” of the first key driver was not part of any scenario, neither was one of the two “network as ‘Achilles’-heel or ‘bottleneck’”-projection of the fifth key driver. For the sixth key driver, the only included projection in all 100 scenarios was the first one (“Overall large increase in sensors”).

The four scenarios were chosen using the criteria described in section 3.2 (pp.24). In Figure 9, the four clusters of these four scenarios are circled in orange and connected with lines to show that this cluster selection can cover a large scenario space. Attachment xyz includes these four scenarios and the corresponding projections. As described, the four scenarios were adapted based on the results of the interview with E10, and condensed into three final scenarios.

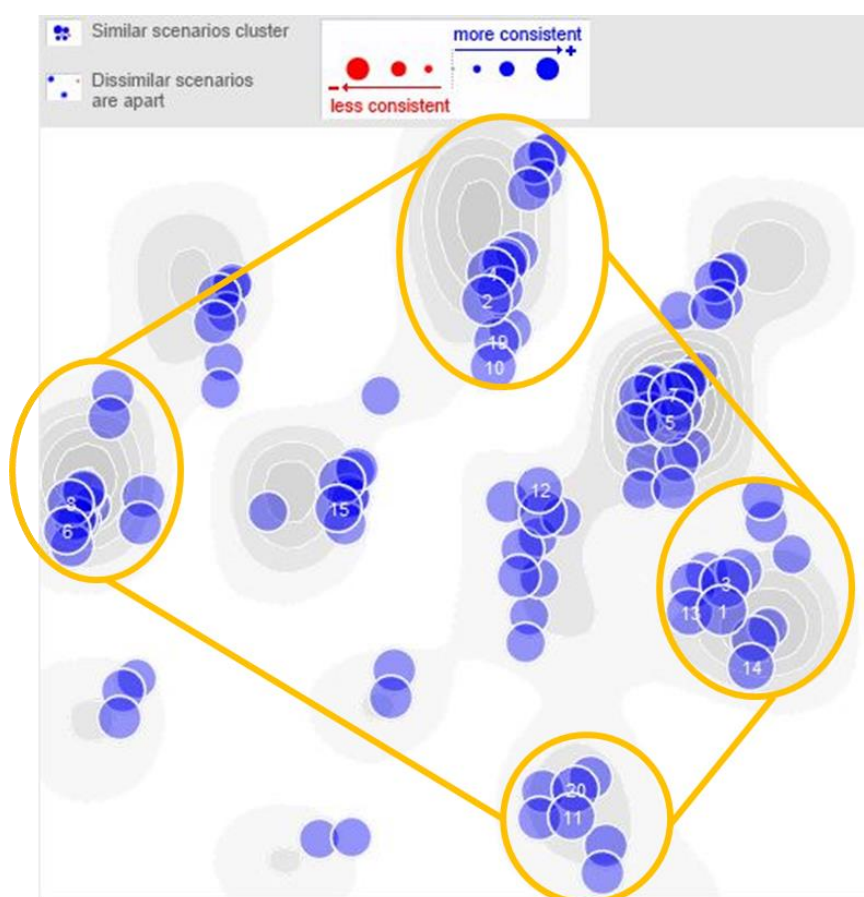


Figure 9: Scenario cluster selection

The three final scenarios for the future development of RIMDB by 2025 are narratively described in the following sections. As already mentioned, the scenarios do not serve to predict the future with a certain probability, but should serve to open the space of possibilities that are considered for strategic decision-making.

4.2.1 Scenario 1: The open-source revolution

Over the years leading up to 2025, the number of devices with embedded sensors has grown by the billions, creating the IoT for practically all environments. This could happen, because businesses across industries acknowledged the enormous benefits of collecting data at various points and heavily invested in installing such sensors. There were also less investment-heavy initiatives supporting the development of the IoT, such as the Phonvert project which triggered people all over the world to recycle their old smartphones to an IoT device which can now collect data²². Moreover, businesses realised that by letting their customers (both private consumers and businesses) reap bottom-line value from the IoT devices and the data these devices collect, customers are much more accepting of the new products and services that ensued. Customers now trust businesses with their data and are more willing to disclose it. They also invest in new IoT products and services that are on offer.

Because data is collected in many ways in the IoT, it is very multi-faceted: Data can originate directly from humans and their behaviour (e.g. sound recordings, video material, written text), as well as from the environment (e.g. weather data, data on air pollution) and from machines (e.g. a smart factory robot which acts autonomously while producing goods). In consequence, different big data analysis systems are necessary to treat these different types of data. There are separate solutions for structured and unstructured data, but also more and more combinatory solutions that integrate systems.

In order to actually gain value from the IoT, businesses need to analyse the data they collect. Real-time response has become the state of the art, and the demand for systems supporting real-time data analysis is strong. Service delivery was accelerated to a real-time standard as well. Those businesses who until now failed to install real-time systems for analysing large data volumes are struggling and starting to fall behind their competition: Not only can those businesses implementing real-time systems make strategic decisions much faster and react directly to (or even predict) customer needs, but real-time systems also make completely new business models possible. For example, car insurances can now personalise their pricing and coverage schemes to the driving behaviour and environmental conditions in real-time by collecting the necessary data via sensors in the car and its environment. Those driving safely and in less accident-prone areas benefit from lower insurance prices. When transporting passengers, the premium increases slightly so that additional costs in the case of an accident are also covered. This dynamic pricing and coverage system offers many benefits for drivers, and the new real-time insurances are able to gain competitive advantage over those that still use the old-fashioned static schemes. Both sides profit: The driver paying car insurance profits from fairer pricing and is rewarded for safe driving. Moreover, he can be certain that all potential damage is covered, e.g. when transporting a number of other people. The car insurance can estimate risks better and thus reduce its losses. Many other examples for real-time products, services and

²² For more information, see <https://phonvert.com/>

business models exist. Especially the first movers are able to capture market share from those that adopt too slowly, because they can establish themselves as price leaders for a certain time.

Generally, the notion of “data is good” has become accepted due to the immense benefits for all sides, and previous fears of “big brother is watching you” have been set aside.

Governments and regulatory institutions have become aware of the large socio-economic benefits these developments can evoke. Openness for the collection, use and sharing of data is supported strongly, and restrictions previously blocking this trend have been limited. For example, the EU’s strategy for a Digital Single Market in the EU, including the proposition of a “European free flow of data initiative” (EC, 2015a), was implemented. The EU-US Privacy Shield agreed upon in February 2016 (EC, 2016a; EC, 2016b) was trimmed down so that data can now move more freely between the US and the EU than ever before. However, governmental institutions also want their piece of the data pie. For example, building on first pilots (Heaton, 2015), more projects were developed in the field of predictive policing to help crime prevention (e.g. by tracking known perpetrators and using algorithms to identify likely areas and times for specific crimes) or help fire fighters with identifying potential hot spots. Having recognized the huge potential of data use and its possibilities, governments exploit data for various other purposes too. For instance, they use data from the industrial sector in order to try to improve the international competitiveness of their own regions in comparison to other regions. So not only is the attitude towards exploitation of data generally open, but there is also competition between various entities for who can reap the most value from the available data. This further highlights the relevance big data analysis systems play for governments in the scenario. In addition, data analyses are used by governments to identify and track down individuals, e.g. who reside in a place without a right of residence or who have a high potential for committing crimes such as tax fraud. Considering these intentions, it is clear that governments are one major customer group for real-time big data analysis systems.

All this would not be possible without sufficient connectivity to transfer the data from the IoT devices to where it is stored and analysed. Therefore, the mobile network and overall connectivity has been boosted over the last years: Large-scale investments by new network and connectivity providers increased competition and gave an incentive for incumbent network providers to invest heavily. For example, the 5G network was developed as planned. Technologies in which newcomer network providers invested include BTLE and the LiFi technology, which further support ubiquitous connectivity.

Thanks to the well-developed connectivity, as well as government and regulatory support, the IoT can generate immense data volumes. That is why the hardware for systems that store and analyse this data had to become more efficient. Despite the fact that hardware became cheaper over the years in general, because of the necessary increase in efficiency, hardware component costs stagnated, which

is why the hardware cost aspect for big data analysis systems (i.e. also RIMDB) stays the same for users.

However, a certain cost benefit for users comes from the fact that the most common RIMDB nowadays are not commercially distributed proprietary RIMDB, but open-source RIMDB: Now, in 2025, they have taken over the market. A few years earlier, in cooperation with Oracle and open calls to the developer community, Google further developed its DB Cloud SQL (a cloud-version of the relational open-source DB MySQL), so that it now runs completely in memory. With that, Google created a well-working open-source RIMDB available through the cloud and to the developer community. The latter continues development, so that it set a very high standard for all RIMDB. Google profits from this move by offering highly targeted additional services to their users. These services are based on further analyses of the data that their users upload to the Cloud SQL DB, and comparing data sets and queries from different users. Naturally, the data of one user is not shared with other users without permission, as this would nullify the value of the data. However, through these add-on services Google can for example suggest new types of analyses to the DB users, or suggest other users as potential partners for sharing data. Considering the incredibly large amounts of data at Google's fingertips, the additional suggestions and offers they can make to their users are widely unmatched. In the wake of these events, a few other open-source RIMDB have emerged and attained a solid user base. To retain some market share, the previously commercial RIMDB vendors felt pressured to open the code of their RIMDB as well, and moved away from business models where RIMDB licenses are a central source of income. However, the new open-source in-memory Cloud SQL DB, and other emerging open-source RIMDB, have cost them a significant part of market share. The previously proprietary RIMDB vendors are struggling to retain some market share with their open-source RIMDB due to the strong position of other open-source RIMDB, most of all the in-memory Cloud SQL.

4.2.2 Scenario 2: The big data golden age

In the second scenario, some aspects are similar to those of the first scenario. Similarly, the IoT has developed to billions of connected devices with embedded sensors, and ubiquitous connectivity was achieved. The IoT thus generates large data volumes ready for analysis. In the same manner as previously explained, real-time response systems have become the golden standard for reaping maximum value from the available data, creating new business models, services and products. As explained, there are a number of big data analysis solutions available because all these data in the IoT come from different sources and are more or less structured.

From government- and regulatory-side, a similar shift towards openness for exploitation of value from data is observable – albeit in this case, socio-economic benefits are the main reason. Also in the second scenario, regulations such as the EU-US Privacy Shield which could potentially handicap business around the collection and use of data were cut in a way that enables free data flows. The EU

has also enforced its Digital Single Market. One example on a national level is the Digitale Strategie 2025 by the German Federal Ministry for Economic Affairs and Energy (BMWi) from March 2016 (BMWi, 2016), which was implemented over the last years. The strategy supported the optical fibre network for fast connectivity, which is now installed nationwide, and efforts for modernisation of Germany as a key location for industry, especially when it comes to bringing Industry 4.0 and digitisation to SMEs. A new Federal Digital Office was also installed to support all matters concerning digitisation and reaping the value from data. Overall, many governments and regulatory institutes thus do their best to support full exploitation of data. Industrial espionage is less of a concern – the central focal point is using data for the value it can provide both socially and economically.

Unlike in the first scenario, both open-source and commercial, proprietary RIMDB solutions exist next to each other. On the one hand, open-source RIMDB caught up to the proprietary solutions, setting a high standard for all. On the other hand, commercial RIMDB vendors now strive to provide additional value with their proprietary RIMDB in order to exceed what open-source solutions can offer. They do this by making specific data bundles available to their customers when acquiring a license for their proprietary RIMDB, offering additional support services for project implementation and providing example projects for various contexts. Moreover, the commercial vendors are both liable and supportive when a system malfunctions or produces errors. The most important user group for open-source RIMDB are SMEs, while larger corporations usually prefer the proprietary RIMDB due to the additional benefits they offer and because they can rely on the vendors to fix errors.

Because of the ubiquitous connectivity now present, cloud solutions have become increasingly popular over the last years. The fact that governments are so supportive of open collection, use and sharing of data also enabled the development and growing popularity of cloud solutions, and previous restrictions for data security and privacy do not hinder cloud usage anymore. For example, EU regulations do not restrict the use of cloud servers located in the USA for European citizen data anymore. That is not to say that data is not protected. The cloud providers strive to affirm their integrity, since the value of their customers' data would be lost if it were not safely stored. The trust in the cloud was further supported through the emergence of open-source RIMDB as an alternative to proprietary RIMDB, because their code can be reviewed openly for potential leaks, giving more control to the user. Overall, cloud usage is now better supported. The economies of scale such cloud solutions offer have decreased the costs for big data analyses on customer side. As a result, it is now possible for many more start-ups and SMEs to use RIMDB solutions for big data analysis (as well as other big data analysis systems). This effect is heightened by the availability of open-source RIMDB as an alternative to proprietary RIMDB. Businesses now have more choice between both open-source and commercially distributed proprietary RIMDB solutions, increasing competition in the market. For the overall economy, considering these developments together with the support of governments and diffusion of cloud systems, many proclaim that 2025 is the golden age of data.

4.4.3 Scenario 3: Platforms for the win

The IoT is just as far developed in scenario three as it is in the previous two scenarios. Ubiquitous connectivity was achieved as well – although the network and connectivity development required a large-scale government investment effort, as private investors felt it would have been too much of a financial burden for them to boost the network and connectivity development far enough. Because of less private competition, there are not that many innovative solutions for achieving ubiquitous connectivity in the market. Nevertheless, openWifi initiatives and a better network coverage through government investment make the IoT and connection of all its devices with embedded sensors possible. Also in this case, initiatives such as the German Digitale Strategie 2025 -amongst others- pushed the development of the optical fibre network forward.

As the investment in the network and connectivity development shows, governments support the trend towards IoT and big data. Therefore, similar to the previous scenarios, regulatory restrictions previously inhibiting open collection, use and sharing of data -also across borders- are now trimmed down so that maximum socio-economic benefits can be reaped from the IoT and the data it makes available.

Because of all this support for the IoT, the data volumes it creates are enormous. Hardware requires higher efficiency than before for dealing with these data volumes. Although prices for hardware decreased in general, the costs for big data solutions have been staying the same for many years when it comes to hardware costs.

However, the cost-aspect did not stop real-time response from becoming the state of the art. Those companies who were not able to adapt and did not implement a real-time system for analysing data have lost the competitive race and were kicked out of the market. Because of their price-sensitivity and the fact that costs for big data analysis solutions (including RIMDB) continue to be rather high, SMEs were particularly affected. For many companies, their products, service offers and business models were not competitive anymore. In this case, initiatives such as the German efforts for supporting Industry 4.0 and digitisation for SMEs (see e.g. BMWI, 2016: Mittelstand 4.0 – Digitale Produktions- und Arbeitsprozesse, Digitalisierungsoffensive Mittelstand) failed to a large extent. However, the fact that some firms had to cease business also opened up room for new digital businesses. Those companies that are still in the market already have implemented a real-time big data system and offer services and products in real-time (or even with predictive characteristics). Overall, improvements in response speed for data analysis are just marginal, meaning that the currently used systems have none to extremely low latencies, rendering further efforts to become “more real-time” unnecessary. Most companies therefore just need very occasional updates on their implemented systems, but do not need to buy new systems or more extensive upgrades all that frequently. Hence, the demand for new real-time systems is much lower now in comparison to previous years: Now, in 2025, market saturation has been reached. The only firms in need of completely new implementation of real-time systems are start-ups.

Until recently, the large commercial RIMDB vendors had a comfortable positioning in the market for big data analysis. Now, this market saturation poses a serious threat for the commercial vendors of proprietary RIMDB. However, one RIMDB vendor recognized this development early enough, and created a new business model: Their previously commercially distributed RIMDB was made available as an open-source solution, and is now used as a vehicle for a platform business model. The RIMDB vendor provides the framework for the core technology (i.e. their RIMDB) for free, and developers can write their own solutions for problems. The vendor starts making money as soon as the developers sell their solutions by receiving a share of the revenue. In essence, this business model is similar to other platform businesses such as the App-Store, or AirBnB, where the platform provider starts profiting as soon as a transaction is made. Additionally, some vendors sell a selection of use cases with packages of solutions and relevant data as add-ons. This gives the RIMDB vendor an excellent competitive position: Both the developers and their customers are customers of the RIMDB platform and are indirectly charged a small fee for the service of the platform when a transaction is made. The developers benefit as they have a very little up-front investment, and the users can profit from the large offer on the platform they can choose from. Value is being created at all ends – for the developers, their customers and for the RIMDB vendor as the platform provider.

5 Conclusions for RIMDB and Industry 4.0

Based on the results of this thesis, this section presents the conclusions. Because the research question was twofold, the conclusions from the scenarios will also be split: Firstly, conclusions for RIMDB will be discussed, and secondly, the conclusions for what these scenarios can mean for Industry 4.0.

This thesis cannot and does not aim to provide a perfect template of strategic steps to success when it comes to real-time big data analysis and Industry 4.0. However, as mentioned at various points, the main goal of this thesis is to broaden the point of view as well as the scope of considered possibilities of strategic decision-makers by illustrating various possible future developments in a narrative, holistic way through the scenarios. Thus, this thesis can serve as an aid when formulating strategies and deciding on next steps for businesses. Overall, decision-makers are advised to watch the future development of the six identified key drivers closely over the next years, in order to be up to date in which direction they develop and be prepared for possible changes. Every entity will have to decide for him- or herself which of the results and lines of thought presented in this thesis are indeed incorporated into the decision-making processes. After all, as Alan Kay put it, “The best way to predict the future is to invent it.” (TED Conferences, LLC, n.d., paragraph 1).

5.1 Developments relevant for RIMDB

The possible future developments relating to RIMDB in this thesis can serve as guidance for RIMDB vendors, but can also be highly relevant for other parties involved in big data analysis, such as providers of alternative or complementary technologies or companies planning on implementing a big data analysis system.

All three scenarios foresee a strong development of the IoT, and in fact all of the 100 most consistent scenarios point in this direction. However, one cannot neglect the fact that data sources in the IoT will be multi-faceted and data will be of unstructured, semi-structured and structured nature. Since RIMDB are primarily suitable for structured data, it is thus important for RIMDB vendors to consider in how far they might be able to either partner with, provide gateways to or integrate systems for unstructured or semi-structured data. First examples going in this direction are Microsoft’s Polybase²³, Oracle’s Big Data SQL²⁴, and the integration of SAP HANA with the Hadoop framework²⁵. There is room to further develop solutions in this direction.

In all scenarios, a well-developed network and corresponding connectivity are important as a highly relevant foundation in the big data environment. Looking at the scenarios, the question remains whether government efforts alone can reach sufficient connectivity, as is the case in scenario three, or whether businesses should support the development themselves, e.g. by investing into innovative

²³ For more information, see <https://msdn.microsoft.com/en-us/library/mt143171.aspx>

²⁴ For more information, see <http://www.oracle.com/us/products/database/big-data-sql/overview/index.html>

²⁵ For more information, see <http://www.sap.com/solution/big-data/software/hadoop/index.html>

connectivity solutions such as LiFi or developing their own network grids in areas with low connectivity. Because of bad connectivity, not even near real-time data transfer can be an option for many businesses. By supporting network development and broadening the areas where real-time data transfer is possible, e.g. by providing infrastructure of their own when necessary, RIMDB vendors could gain new customer groups. Otherwise, although in none of the 100 most consistent scenarios this was the case, the network and lack of sufficient connectivity could indeed become a bottleneck for the IoT development, which would mean less valuable data could be collected. This in turn could have negative effects for the diffusion of RIMDB (and other big data solutions).

For RIMDB users, and users of other real-time data analysis systems, the example of insurance providers described in scenario one is very relevant: Competitive advantage from implementation of real-time systems is likely going to be subject to considerable volatility and ephemerality. First movers implementing such systems and creating new business models, products and services may be able to capture considerable market share, but the question remains whether this will be enough to build a sustainable competitive advantage. Businesses should give serious thought to this aspect when working on real-time based business models, products and services. Also, RIMDB vendors (and other big data analysis providers) can also start to consider how they can continue adding value to users once real-time becomes the state of the art.

Along these lines, it may also be interesting for RIMDB vendors to think about the proposed platform business model of scenario three, which could ensure long-term recurring profits even when the market for RIMDB systems is already saturated, or open-source solutions steal a portion of the market share of commercial RIMDB vendors. Naturally, the story of Google and Oracle creating a largely popular open-source RIMDB based in the cloud of scenario one is only an example based on current developments and trends, and should serve to show that it is not impossible that proprietary RIMDB may lose their positioning. Such a platform business model, as proposed in scenario three, may therefore be more sustainable on the long run when competing with big data giants such as Google or Facebook. As Van Alstyne of the MIT Initiative on the Digital Economy put it, “Platforms beat products every time.” (Accenture, 2016, p.42). Overall, RIMDB vendors are advised to consider the possibility that open-source RIMDB –and for that matter, also other open-source big data analysis solutions– could play an increasingly relevant role and gain market share.

Lastly, in particular the first scenario shows that governments may also be a highly relevant customer group for big data analysis systems, including RIMDB. Partnering with governments may also be interesting when it comes to open data which could be pre-analysed or provided as add-on packages to RIMDB, as described previously.

A certain awareness may need to be raised in the more price-sensitive SMEs with low technological abilities about current digitisation and IoT developments, and about the dangers of falling behind from a competitive perspective. Here, the price point of RIMDBs may be highly relevant. Another important

aspect, which did not appear in the key drivers but is part of the 31 drivers is a lack of sufficient skills in many companies, especially SMEs to handle the new big data technologies. RIMDB vendors (and other providers in the field of big data) could support these companies, firstly in developing these necessary skills, and secondly by making their systems available at more affordable rates. On the long run, such efforts could open up a large new customer group for RIMDB vendors.

On top of that, RIMDB vendors could also consider adding further value for their customers by e.g. selling additional data packages. It may be particularly interesting for companies to mix (structured) transactional data with (unstructured) sensor data. One example case could be combining unstructured data from the sensors of a production robot with (structured) ERP planning numbers to gain further insight, and/or update both systems with real-time changes. This example already points to the next section on conclusions for Industry 4.0.

5.2 Conclusions for Industry 4.0

This section explores what the results of this thesis mean for Industry 4.0, and refers to the four key components of Industry 4.0: (1) Cyber-Physical Systems (CPS); (2) the IoT; (3) the Internet of Services (IoS); and (4) Smart Factories (Hermann et al., 2015). As mentioned, the conclusions in part make use of input from the last expert interview round. These conclusions for Industry 4.0 are mainly relevant for strategic decision-makers in the industrial sector, but may also be interesting for other entities such as governments due to the interdisciplinarity of Industry 4.0.

In the sense that decisions will be made based on data, and automated to a certain extent, E10 is convinced that any industry or business field can be digitised – even in very “human-centric” areas such as medical and geriatric care. Diffusion of real-time systems, as presented in all three scenarios, means that these data-based decisions can be made in real-time. The same will be true for CPS as one aspect of Industry 4.0: In industrial environments, systems combining physical and digital elements will collect data, which will then be stored and analysed. These analyses’ results can eventually be used for various further actions, such as the example mentioned in the previous section: Combining sensor data from a production robot with additional data from e.g. planning systems can optimise both the robot’s production processes and the planning data. This cooperation of CPS also relates to driver 10 “Advances Machine Learning / Deep Learning” Table 4 (p.40-42): For many areas, human interaction will not even be required anymore if the CPS are smart enough and have the data available that they need to fulfil their tasks properly. As Hermann et al. (2015) state, the cooperation of various CPS in real-time via the IoT in an industrial environment will create the mentioned Smart Factories. For operations on a general level, this means that both the process as well as the activities converting resources into goods and services changes. It also means that the characteristics of what is an input, i.e. resource, and what is an output, i.e. good or service, change. This relates to Koch et al.’s (2014) definition of Industry 4.0: In the Industry 4.0, all the data relevant to one activity in the transformation process for creating a good or service will be available in real-time, and connect all people, things and

systems involved in the value chain – “(...) within and across companies.” (p.10). Again, to make this possible, the relevance of further diffusion of real-time systems and all driving factors connected to this development is high.

Similar to what the scenarios show, sufficient connectivity is essential for the formation of Smart Factories. CPS may not all be located in one area, but dispersed all over the world at various production facilities. To make real-time cooperation of CPS possible, the network transporting data from one location to another needs to be adequately developed and fast. This implies a very large scale infrastructural effort. As mentioned in the previous section, the question remains whether government efforts will be sufficient to reach such ubiquitous connectivity. Industrial players should certainly consider the high relevance of network and connectivity, and collaborate with governments and other industrial players to ensure that the network and connectivity development progresses adequately to enable the IoT. One topic that was already mentioned in driver 17 (Table 4, p.40-42) “Industry Standard / Monopolization” is that the IoT needs certain standards, both technologically (e.g. regarding gateways), and on a regulatory level (e.g. regarding security). Because the IoT is part of Industry 4.0, standards are also highly relevant for industrial players. Especially when it comes to regulatory aspects, cooperation of countries and regions, and also with the industrial players, should ensure that standards are sufficiently defined in order to make the IoT work. If IoT standards are not effectual across companies and country borders or regions, it could at least be more difficult (if not impossible) to make the vision of Industry 4.0 a reality as described by Koch et al. (2014), because not all relevant data for one activity will be available in real-time during business operations.

Moreover, industrial players will need to keep in mind that the data they generate, collect and eventually analyse through CPS, and generally the IoT, can come from various sources, as mentioned. Further development of big data analysis solutions that are capable of handling any type of data would thus be in their best interest.

The third scenario, and to an extent also the second one as well, highlight that SMEs may lag behind when it comes to real-time big data analysis and installing new systems, amongst others RIMDB, which can affect their competitive position negatively. As mentioned, likely causes are lack of (1) awareness of the risks of falling behind, (2) of skills to implement these new technologies effectively once awareness has been achieved, and (3) of funds to acquire the necessary technologies. The second cause relates strongly to driver 13: “Workforce Skills / Digital Dexterity” (Table 4, p.40-42), which already describes the problem of lack of skills by the workforce. E10 even states that, due to the developments in Industry 4.0, craftsmanship and skilled labour in production contexts will become entirely unnecessary. From a broader economic viewpoint this could mean a high growth in unemployment in these groups of workers. Both governments and industrial players will need to collaborate to find solutions for this issue, including changes in the education system. Coming back to SMEs, governments, as well as businesses relying on the work of SMEs (e.g. as component suppliers),

need to consider strategies for tackling the potential causes if they want to mitigate the potential negative effects of many SMEs lagging behind technologically. As an example, due to the particular economic relevance of SMEs in Germany, the country's Federal Ministry for Economic Affairs and Energy has already acknowledged this risk, and plans to actively support SMEs and craftsmen in their digitisation efforts and joining the Industry 4.0 with its Digitale Strategie 2025 (BMWi, 2016).

As portrayed in scenario two, open-source RIMDB could be a solution for SMEs to decrease their costs, and thus make it more feasible for them to adopt real-time big data analysis. On the other hand, an open-source system also requires IT skills: The code may need adaptation to the specific context, and unlike proprietary software, open-source solutions do commonly not offer any customer service. Thus, simply pushing open-source RIMDB to SMEs cannot be the only and final solution. The necessary skills to handle the technology, and/or funds for hiring experts will still be required.

Although this section and the previous one present highly interesting considerations for RIMDB and Industry 4.0, this thesis needs to be reviewed under consideration of its limitations, but also the efforts that were undertaken to mitigate them, which are discussed in the following section.

6 Discussion of limitations and future directions

It can be argued that the search and review of literature in those sections of less systematic approach in this thesis could have been conducted in a more structured and systematic fashion. However, for some sections a less systematic approach was deemed necessary given the newness as well as multidisciplinary of the topics. Also, as explained, even when following a more narrative approach, certain selection criteria were considered to provide a certain structure. Given time and resource constraints, a completely exhaustive literature review is not possible in the context of a master thesis. When conducting qualitative interviews, as done for this thesis, Saunders et al. (2009) note that certain data quality issues can occur: Response bias may ensue from verbal or non-verbal signals from the interviewer as well as her tone of voice, or the topic under discussion, thereby distorting the interviewee's response in one direction or another. Also, bias in the interviewee's responses may have resulted from their personal or professional background. The interviewer tried to avoid such bias as best as possible, but cannot be certain this was achieved. The criticism of qualitative interviews potentially having reduced generalizability (amongst others due to small sample size and lack of standardization in the interviews) does not apply in this case as the goal of this thesis is not to collect information about the experiences or behaviour of individuals or groups, or any other issue that may require generalization. Rather, this thesis aimed at collecting information on a specific technological topic and both subjective and objective knowledge from a selection of experts. It is possible that a larger sample size could have broadened the scope in responses, though, which could have resulted in even more valuable insights. Nevertheless, within the means of a master thesis, adequate sample diversity was achieved.

From a methodological perspective, it is additionally worth noting that the principle of the consistency matrix appears not to be very intuitive and therefore rather difficult to comprehend fully for the experts who did not evaluate it face-to-face based on the written instruction sent to them via e-mail. That is why the highest relevance was given to the face-to-face interview for evaluating the consistency matrix. Apart from these aspects, the scenario method and its implementation in this thesis (including the expert sample) were already discussed critically in section 3.2.3 (pp.29).

As mentioned, some projections did not appear in the 100 most consistent scenarios which reduced the scenario differentiation. However, since the scenarios are based on the consistency matrix, and the consistency matrix was evaluated using expert feedback, the researcher should assume that the 100 most consistent scenarios assembled by EIDOS are indeed those scenarios that decision-makers should consider most strongly.

In particular regarding the aspect of Industry 4.0, it was notable that the conclusions did not only relate directly to the scenarios, but also to some of the drivers that were not determined as key drivers and did thus not appear directly in any of the scenarios. However, not only the key drivers are relevant results of this thesis. Moreover, there are many interconnections and relations between the

drivers, as displayed in the driver grid (see fig. xyz), which means that conclusions from one driver may lead to another driver. By applying the principles of a TF study, the 31 drivers were determined mainly for RIMDB. Naturally there were certain overlaps with Industry 4.0 as the topic is so closely interconnected with real-time big data analysis. Nevertheless, it is important to understand that, for the same reason, the scenarios were created specifically for the technology RIMDB. Of course, due to the overlaps of the topics and the formulation of the research question, Industry 4.0 was not completely disregarded in the process. The researcher had initially planned to gather more in-depth insights from the last round of expert interviews in order to bridge between the scenarios for RIMDB and the topic of Industry 4.0, referring to the definition of Industry 4.0 by Koch et al. and the four key components identified by Hermann et al., 2015. This was only achieved to a certain extent – for the most part, because for R4 only one expert could be interviewed (a) due to unavailability of other experts, (b) due to the unwillingness of other experts to participate, and (c) because time and resource constraints did not allow further broadening of the expert search. The one interview conducted did not yield insights to the depth and extent the researcher had initially expected. This could not have been foreseen. It was still possible to identify interesting insights and conclusions for Industry 4.0. However, future research could expand on the aspect of Industry 4.0 in relation to real-time big data analysis. One possibility could be conducting a value chain analysis using a case study approach in order to identify specific activities in which real-time data could add further value or might have a particularly transformative effect to the activity. This information could help decision-makers in identifying those areas in their operations on which to focus especially during the transformation towards Industry 4.0, e.g. regarding where to place sensors at existing factory machinery, where to link activities producing symbiotic data and which workers to train with new skills.

For Industry 4.0 another aspect described by E10 could be interesting for future research: CPS in the sense of merging of humans and machines. One example could be a prosthetic arm connected to both a human brain and a real-time big data analysis system, thus integrating information from external sources in real-time in order to guide decision-making of the human. This may sound very futuristic, but the so called “mind-interface” is indeed a trend that has been recognized and is being researched (see EIT Digital IVZW, 2016). For business operations this could mean that in the Industry 4.0, some activities will require much closer interaction of humans and machines than imaginable today. Latency-free real-time data transfer combining various sources would certainly be a necessary basis for CPS merging a human and machine in a much closer way than currently. For RIMDB vendors, and other big data analysis system providers, it could also be interesting to assess how their soft- and hardware may be able to integrate with such futuristic avenues. For decision-makers, it could be very important to identify those areas where such systems could truly add value, similar to what was explained above.

7 Acknowledgements

I hereby like to express my gratitude to my supervisors Prof. Dr. Katharina Hölzle and Victoria Götz, M.Sc. for their engagement, constructive guidance and advice during the entire research and writing process of my thesis. I also like state my appreciation to my second supervisor Dr. Michel Ehrenhard for his useful suggestions and feedback throughout the steps towards completing the thesis.

Moreover, I like to sincerely thank the experts for sharing their precious time and valuable insights during the interviews and e-mail correspondence.

Finally, I would like to voice my thankfulness to my family and friends, who have been very supportive throughout the entire process by lending me their ear during difficult times as well as keeping me motivated and inspired.

Julia Reichelt

Appendix

Attachment A: Interview Notes Phase R1	68
Interview Notes E1	69
Interview Notes E2	75
Interview Notes E3	79
Interview Notes E9	83
Attachment B: Interview Transcripts Phase R2	84
Interview Guidline Phase R2	85
Transcript E4.....	87
Transcript E5.....	96
Transcript E6, E7	101
Transcript E8.....	112
Attachment C: Interview Transcripts and E-Mails Phase R3.....	120
Fact Sheet for Phase R3	120
Transcript E1.....	121
E-Mail E4	130
E-Mail E6	132
Attachment D: Interview Transcript Phase R4	133
Transcript E10.....	133
Attachment E: Parmenides EIDOS TM output files	140
Grid of 31 Drivers	140
Active-Passive Map of Drivers	141
Initial Scenario Choice	142
Final Scenario Choice: Rough Draft Projections.....	144

Attachment A: Interview Notes Phase R1

Interview Notes E1

Questions on databases using in-memory technology		
#	QUESTION	ANSWER
1	Which databases using in-memory technology are you aware of? Welche IM Datenbanken kennen Sie?	Redis (open-source) Hazelcast SAP HANA Von IMB und oracle: kennt er, aber hatte damit nicht zu tun bisher
2	Do you see any specific limitations for the development and diffusion of in-memory technology, and more specifically in-memory databases? How would you assess the relevance of these possible limitations? Sehen Sie mögliche Einschränkungen für die Verbreitung von IMD auf dem Markt? Für wie relevant halten Sie diese Einschränkungen? Hoch, mittel, niedrig, sehr niedrig Scale: - High - Medium - Low - very low Please explain your answer briefly	Positive: Früher war storage so teuer: preis sinkt immer weiter → preisverfall: memory hat sich gut entwickelt, wird sich auch noch weiter entwickeln WEITERER PREISVERFALL fördert entwicklung Negativ: Rein technologisch: thema big data: man versucht mit normaler hardware (SAP HANA läuft ja auch nur auf appliance die gewisse hardware erfüllen muss) → hadoop und spark: versuchen mit normaler hardware große datenmengen zu verarbeiten → keine hardware appliances wie bei SAP HANA NÖTIG thema big data spielt gegen entwicklung von SAP HANA selbst SAP sagt HANA ist nicht mehr für jeden case zu teuer – nicht für alle cases geeignet aus dem aktuellen markt nehmen sie war: spark, hadoop und big data wachsen stärker

		SAP HANA sieht er weniger als „big data“ thema: Versuchen aus sensoren infos zu sammeln (oder aus anderen systemen): SAP HANA: thema analytic haben die sich auf die flagge geschrieben – rüsten gerade auf, im rahmen big data wird hadoop/spark stärker wachsen im vordergrund, aber SAP HANA wird im bereich stärker wachsen wo ich meine normalen business anwendungen hab (e.g. ERP3 wird auf SAP HANA umgestellt) Für normale SAP anwendungen (wie CRM) als unterliegende technologie um analysen zu ziehen ohne big data cluster aufzumachen Normale SAP Suite – wird weiter wachsen im bereich Big data: daten sammeln aus anderen geräten werden hadoop und spark (streaming) mehr wachsen
3	How would you assess the degree of diffusion of in-memory technology, and again more specifically in-memory databases today? How will things look in 2020 and in 2025 according to your opinion? Für wie hoch halten Sie den Verbreitungsgrad von IM? Sehr hohe Verbreitung, hoch, mittel, sehr niedrig Inwiefern könnte sich dies bis 2020 und bis 2025 verändern? INFO: Der Verbreitungsgrad von IM wird anhand der Anzahl von Akteueren, Organisationen und Instituten, die sich mit dem technischen "know-how", Wissen und Erfahrungen zu dieser Technologie beschäftigen bewertet. Scale: - very high - high - medium - very low Please explain your answer briefly [NOTE: In this case, the degree of diffusion of in-memory technology/databases is	Kennt dazu keine statistiken: IMD in relation zu traditionellen datenbanken: IMD im verhältnis aktuell noch gering wird bis 2020 auch durch den preisverfall und dadurch das SAP das thema auch puscht steigt in richtung mittel, und wenn der preisverfall weiter sich forsetzt wird es 2025 schon einen hohen verbreitungsgrad haben Aber weniger für thema big data, sondern business andwendungen (normale datenbank wie oracle wird durch IMD ausgetascht) – SAP setzts sich dadurch

	assessed by the number of actors, organisations and institutes with specific technological know-how, knowledge and/or experience for this technology.)	
4	Which fields of application can you imagine for in-memory technology, and specifically in-memory databases? Do you see certain limitations and for the application of this technology? Welche Anwendungsfelder können Sie sich für IM/IMD vorstellen? Sehen Sie hier bestimmte Limitationen/Einschränkungen für die Anwendung der Technologie? <i>Please assess how many application fields you see and how large these fields are. Name a maximum of twenty application fields specifically and choose those you consider most relevant (if more, please only number of fields).</i>	Business suite: CRM, ERP, business anwendungen (lässt sich auch durch das wo SAP sich aktuell bewegt untermauern – ich muss nicht mehr e.g. warten auf umsätze wenn ich eine analyse will) Auch anwendung für big data, aber weniger für hadoop und spark Momentan sieht SAP anwendung bei IoT und M2M, aber er denkt dass da das thema spark stärker wachsen wird (es gibt aber schon anwendungsfälle) Da ergeben sich auch use cases durch das schnelle verarbeiten von sensor informationen, es soll ja mal alles vernetzt werden Wo er sich das gar nicht vorstellen kann: für klassische intranet-anwendungen im kleinerem umfang die irgendwelche daten administrieren SMEs: Sieht schon anwendungsfälle, eher im bereich für business anwendungen – technologie wird günstiger, grade vllt noch nicht richtiger zeitpunkt, aber wird noch kommen HANA ist ja auch cloud-basiert, d.h. die SMEler müssen nicht hardware/appliances kaufen, d.h. das wird günstiger (auch für business anwendungen weil die aus der cloud heraus beziehen kann) Schon auch geeignet Cloud spielt für SMEler eine größere rolle
5	What are potential customer groups for in-memory technology, and specifically in-memory databases? <i>Please explain briefly why you consider this customer group relevant</i>	Großunternehmen als hauptplayer (dort gibt es viele daten aus verschiedenen bereichen die man in real-time auswerten will) Mittelstandsunternehmen eher mit cloud-lösung Kleinere unternehmen: IMD ist zu teuer, und bei kleineren datenmengen muss ich

		in memory verarbeiten ■ über alle bereiche hinweg: pharma, automotive, jedenfalls was BI angeht haben alle die gleichen anforderungen ■ klassische produktionsunternehmen: haben auch CRM und ERP anwendungen – auch relevant ■ wo es interessant wird bei industrie 4.0: real-time daten mit meiner fabrik (robotik etc.) verbinden, teilweise auch auf losgröße eins gehen – aber eher thema spark (streaming)
6	Please make an educated guess on the intensity of research and development and the type of development. Können Sie abschätzen wie intensive Forschung & Entwicklung für IM/IMD betrieben wird? Die Intensitätsgrade sind: - mittel (Grundlagenforschung) - hoch (Anwendungsfelder) - niedrig (Kostenreduktion) - sehr niedrig (reife Technologie) Scale: - Medium (basics) - High (applications) - Low (cost reduction) - Very low Please explain your answer briefly	Glaubt schon, dass IM technologisch sehr weit, erwartet da nicht so viel neues (hat sich über die jahre was entwickelt), SAP hat ja wohl auch schon von Oracle was abgucken) Glaubt dass sich da forschungsmäßig nicht so viel entwickelt Hauptforschung ist abgeschlossen, da wird keiner mehr auf die idee kommen IMD von neuem aufzuziehen Wird sich künftig eher differenzieren in richtung - Preis - Verbreitung - Benutzbarkeit (wie einfach lässt sich eine applikation anwenden)
7	Please assess how long you think it will take for the technology to reach the stage of maturity, or whether it may already be mature. Wie lange wird es Ihrer Meinung nach brauchen bis die Technologie reif ist (d.h. weniger Forschung nötig, Anwendungsfelder bekannt etc.): Scale: - Long term (7-15 years until maturity) - Medium term (2-4 years until maturity)	Reife technologie

	- <i>Short term (>1 year until maturity)</i> - <i>Already mature</i> <i>Please explain your answer briefly</i>	
8	Please assess in how far in-memory technology, specifically in-memory databases are available for potential customers. If you think availability differs in certain regions (e.g. Europe, USA, Southeast-Asia, etc.), please specify. Inwiefern sind IMD schon für potenzielle Kunden verfügbar? - Sehr eingeschränkt verfügbar (e.g. nur ausgewählte Test-Kunden) - Es gibt Wartezeiten für potenzielle Kunden (z.b. invite-only beta) - Marktorientierte Verfügbarkeit (gezielt für best. Kunden verfügbar) - Massenproduktion (für jeden verfügbar) Gibt es regionale Unterschiede? (Europa VS USA vs Asien etc.) Scale: - <i>Very limited availability</i> - <i>Waiting periods for potential customers</i> - <i>Availability is market oriented</i> - <i>Mass production</i>	Wüsste er jetzt nicht (regionale Unterschiede) Software läuft auf hardware aspekt: verfügbarkeit eingeschränkt über preis
9	What are your thoughts on the business potential databases using in-memory technology? Wie schätzen Sie das Geschäftspotenzial von IM ein?	-
10	Please describe the influencing factors and potential drivers you consider relevant for the future development of databases using in-memory technology. Welche Faktoren könnten die zukünftige Entwicklung von IM beeinflussen? Relevant könnten sein: - Marktentwicklungen (auch: Konkurrenz!) - Verbreitung der Technologie - Technologische Komponenten und Aspekte der Technologie	-

d) Etc.

Relevant for the future development of this technology are factors and drivers concerning

a) market development

b) diffusion of the technology

c) technological aspects

d) and any other potential aspects you may think of

Please feel free to name several influencing factors and drivers, even if you are not entirely sure.

Interview Notes: E2

Questions on databases using in-memory technology		
#	QUESTION	ANSWER
1	Which databases using in-memory technology are you aware of? Welche IM Datenbanken kennen Sie?	Oracle TimesTen SAP HANA VoltDB
2	Do you see any specific limitations for the development and diffusion of in-memory technology, and more specifically in-memory databases? How would you assess the relevance of these possible limitations? Sehen Sie mögliche Einschränkungen für die Verbreitung von IMD auf dem Markt? Für wie relevant halten Sie diese Einschränkungen? Hoch, mittel, niedrig, sehr niedrig Scale: - High - Medium - Low - very low Please explain your answer briefly	Das Problem sind eher die Software-Stacks, die darauf aufbauen - kaum eine Enterprise-Anwendung ist darauf ausgelegt, In-Memory-Datenbanken in ihrem vollen Umfang zu nutzen. Die gewonnene Geschwindigkeit wird z.B. durch Legacy-Software zunichte gemacht oder die Daten in eine Excel-Tabelle geholt, weil es an besseren und verständlicheren Schnittstellen zur Datenbank mangelt. (Letzteres zu beheben ist u.a. ein kleines Forschungsthema bei uns am Fachgebiet. Dafür suchen wir auch mögliche Anwendungen, um realistische Ergebnisse zu erhalten.) Eine Idee ist, nicht die Datenbank oder Technologie "alleinstehend" zu vermarkten, sondern in Verbindung mit einer Plattform (z.B. HCP, die Hana Cloud Plattform). Kunden werden hier eigentlich kaum etwas von der Datenbank mitbekommen, außer dass Reporting, Forecasting und andere Prozesse deutlich schneller sind. Geschäftslogik kann man dann in Form von Apps aus einer Art App-Store hinzufügen. Auch zu Frage 2: Neben dem Aspekt der mangelnden Awareness/Mentalität haben einige Unternehmen schlichtweg zu viele Daten (bzw. ist Hauptspeicher noch immer teuer) und müssen vorsichtig auswählen, welche sie in-memory verwalten und welche an klassische DBs auslagern.
3	How would you assess the degree of diffusion of in-memory technology, and again more specifically in-memory databases today? How will things look in 2020 and in 2025 according to your opinion? Für wie hoch halten Sie den Verbreitungsgrad von IM? Sehr hohe Verbreitung, hoch, mittel, sehr niedrig Inwiefern könnte sich dies bis 2020 und bis 2025 verändern?	-

	INFO: Der Verbreitungsgrad von IM wird anhand der Anzahl von Akteuren, Organisationen und Instituten, die sich mit dem technischen "know-how", Wissen und Erfahrungen zu dieser Technologie beschäftigen bewertet. <i>Scale:</i> - very high - high - medium - very low <i>Please explain your answer briefly</i> [NOTE: In this case, the degree of diffusion of in-memory technology/databases is assessed by the number of actors, organisations and institutes with specific technological know-how, knowledge and/or experience for this technology.]	
4	Which fields of application can you imagine for in-memory technology, and specifically in-memory databases? Do you see certain limitations and for the application of this technology? <i>Welche Anwendungsfelder können Sie sich für IM/IMD vorstellen?</i> <i>Sehen Sie hier bestimmte Limitationen/Einschränkungen für die Anwendung der Technologie?</i> <i>Please assess how many application fields you see and how large these fields are. Name a maximum of twenty application fields specifically and choose those you consider most relevant (if more, please only number of fields).</i>	Was konkrete Anwendungen angeht, hatte ich mit Victoria schonmal besprochen; kurzgefasst sind das Reporting / Business Analytics, aber auch klassische Data-Mining und Machine-Learning-Probleme wie Forecasting, Text-Mining/NLP, medizinische Studienauswertung, Genomanalyse, Geodaten-Analyse, Targeting/Tracking von Nutzern und Werbung, Empfehlungssysteme, Suchmaschinen, Verkehrsplanung/-überwachung etc. Im Prinzip alles, wo man den gesamten Datenbestand als "Input" für eine einfache Anfrage nutzt. "Neu" ist davon eigentlich nichts, nur gewisse Sachen gehen schneller.
5	What are potential customer groups for in-memory technology, and specifically in-memory databases? <i>Please explain briefly why you consider this customer group relevant</i>	-
6	Please make an educated guess on the intensity of research and development and	Aus meiner Perspektive herrscht gerade große Bereitschaft, R&D

	the type of development. Können Sie abschätzen wie intensive Forschung & Entwicklung für IM/IMD betrieben wird? Die Intensivitätsgrade sind: - mittel (Grundlagenforschung) - hoch (Anwendungsfelder) - niedrig (Kostenreduktion) - sehr niedrig (reife Technologie) Scale: - Medium (basics) - High (applications) - Low (cost reduction) - Very low Please explain your answer briefly	finanziell zu unterstützen - gerade weil es an neuen Anwendungen mangelt. Viele Labs (z.B. Oracle Labs oder SAP ICP) sind mit dieser Technologie beschäftigt... und wenn ich zurückverfolge, worüber meine Doktorandenstelle finanziert ist, lande ich im Endeffekt auch bei einem großen In-Memory-Datenbankanbieter, für den wir gelegentlich Probleme lösen ;)
7	Please assess how long you think it will take for the technology to reach the stage of maturity, or whether it may already be mature. Wie lange wird es Ihrer Meinung nach brauchen bis die Technologie reif ist (d.h. weniger Forschung nötig, Anwendungsfelder bekannt etc.): Scale: - Long term (7-15 years until maturity) - Medium term (2-4 years until maturity) - Short term (>1 year until maturity) - Already mature Please explain your answer briefly	ich sehe den technologischen Lebenszyklus vermutlich eher aus der Zukunft. Von diesem Standpunkt aus würde ich die Technologie an sich sogar schon als ausgereift bezeichnen (Frage 7). Es gibt einige Datenbanken, die sie implementieren (zu Frage 1: Oracle TimesTen, SAP HANA, VoltDB) und auch von wissenschaftlicher Seite ist das alles gut verstanden.
8	Please assess in how far in-memory technology, specifically in-memory databases are available for potential customers. If you think availability differs in certain regions (e.g. Europe, USA, Southeast-Asia, etc.), please specify. Inwiefern sind IMD schon für potenzielle Kunden verfügbar? - Sehr eingeschränkt verfügbar (e.g. nur ausgewählte Test-Kunden)	-

	- Es gibt Wartezeiten für potenzielle Kunden (z.b. invite-only beta) - Marktorientierte Verfügbarkeit (gezielt für best. Kunden verfügbar) - Massenproduktion (für jeden verfügbar) Gibt es regionale Unterschiede? (Europa VS USA vs Asien etc.) Scale: - <i>Very limited availability</i> - <i>Waiting periods for potential customers</i> - <i>Availability is market oriented</i> - <i>Mass production</i>	
9	What are your thoughts on the business potential databases using in-memory technology? Wie schätzen Sie das Geschäftspotenzial von IM ein?	-
10	Please describe the influencing factors and potential drivers you consider relevant for the future development of databases using in-memory technology. Welche Faktoren könnten die zukünftige Entwicklung von IM beeinflussen? Relevant könnten sein: e) Marktentwicklungen (auch: Konkurrenz!) f) Verbreitung der Technologie g) Technologische Komponenten und Aspekte der Technologie h) Etc. Relevant for the future development of this technology are factors and drivers concerning a) market development b) diffusion of the technology c) technological aspects d) and any other potential aspects you may think of <i>Please feel free to name several influencing factors and drivers, even if you are not entirely sure.</i>	-

Interview Notes E3

Questions on databases using in-memory technology		
#	QUESTION	ANSWER
1	Which databases using in-memory technology are you aware of? Welche relationalen IM Datenbanken kennen Sie?	SAP HANA; Exasol (Gfk; Otto) --> persönlich fast erfolgreicher als HANA; dann gibt es Apache Apax Konkurrenztechnologien: - Ganz klar Sachen die nicht ausschließlich auf in-memory arbeiten - Sowas wie Spark, Hive, Spark SQL und Impala (?), bevor sie schauen ob sich in-memory DB lohnt - Meiste kunden haben hadoop cluster und nutzen was damit geht - Apache Drill - Diese technologien sind nur teil-in-memory; die nutzen den speicher nur zur join-optomierung, aber nur wegen des lebenszyklus eines einzelnen query - HANA und co haben komplette datenspeicherung in-memory - Ursprungsdaten in Konkurrenztechnologien Seine kunden haben meistens 20-30-40-50 terabyte bis in tetabyte bereich - Da ist HANA einfach nicht bezahlbar, da gibt es keine lizenz mit vernünftigen preis - Ist sehr begrenzt auf kleine datenmengen → viele kunden brauchen doch andere technologien, die auf der festplatte liegen; genau dafür ist impala zuständig → liest von der festplatte, aber trotzdem sehr schnell
2	Do you see any specific limitations for the development and diffusion of in-memory technology, and more specifically in-memory databases? How would you asses the relevance of these possible limitations? Sehen Sie mögliche Einschränkungen für die Verbreitung von relationalen IMDB auf dem Markt? Für wie relevant halten Sie diese Einschränkungen? Hoch, mittel, niedrig, sehr niedrig	- HANA und exasol → kosten als größter einschränkungsfaktor - Grundsätzlich in-memory DB: grundsätzlich ist das ne gute sache, aber es gibt noch von keinem in der big data welt von einem der größten - Cloudera, Horton-works, mapr → drei größte big data anbieter der welt: keiner bietet eine in-memory datenbank mit an, und 99% der firmen mit big data nutzen eine der drei komponenten → schwierig, wenn man sich das system kauft und dann keinen support für eine der in-memory DB bekommt → kein führer auf dem in-memory open-Source DB markt hat sich noch nicht etabliert

	Scale: - High - Medium - Low - very low Please explain your answer briefly	- Da gibt es auf jeden fall eine chance, dass sich das ändert → dauert aber sicher noch einige Jahre - In-memory DB sind nur für begrenzten einsatzzweck überhaupt nötig, braucht man nur wenn man auf die extrem niedrigen latenzen angewiesen ist → BI analyst ist egal ob er 10s oder 5ms braucht ist egal wenn man sich dafür horrende lizenzkosten spart - Impala und co werden auch immer schneller, aber es wird immer mehr kunden mit diesen geringen latenzanforduernen geben - Ist aber über-gehyped - Da wird es irgendwoe auch open-source variatnen geben, die sich iin den systemen der drei großen wiederfinden
3	How would you assess the degree of diffusion of in-memory technology, and again more specifically in-memory databases today? How will things look in 2020 and in 2025 according to your opinion? Für wie hoch halten Sie den Verbreitungsgrad von relationalen IMDB? Sehr hohe Verbreitung, hoch, mittel, sehr niedrig Inwiefern könnte sich dies bis 2020 und bis 2025 verändern? INFO: Der Verbreitungsgrad von IM wird anhand der Anzahl von Akteueren, Organisationen und Instituten, die sich mit dem technischen "know-how", Wissen und Erfahrungen zu dieser Technologie beschäftigen bewertet. Scale: - very high - high - medium - very low Please explain your answer briefly [NOTE: In this case, the degree of diffusion of in-memory technology/databases is assessed by the number of actors, organisations and institutes with specific technological know-how, knowledge and/or experience for this technology.]	Von seinen kunden 90% nutzen es nicht → er ist Hadoop-berater, seine kunden haben große Hadoop-Cluster Er hat 40 verschiedene kunden, davon hatten 4 exasol oder HANA = 10% (2 davon haben IMDB wieder abgeschafft weil zu teuer und nicht praktikabel) Ist natürlich nur ein ausschnitt; andere technologien sind bei weitem aber mehr verberietet Bis auf einen kunden, alle kunden die er kennt, auch von kollegen, probieren HANA wieder los zu werden, weil es so teuer ist, und weil andere technologien immer besser werden Das ist aber nur ein ausschnitt, insgesamt geht es sicher nach oben Grad HANA ist unverschämt teuer

4	Which fields of application can you imagine for in-memory technology, and specifically in-memory databases? Do you see certain limitations and for the application of this technology? Welche Anwendungsfelder können Sie sich für relationale IMDB vorstellen? Sehen Sie hier bestimmte Limitationen/Einschränkungen für die Anwendung der Technologie? <i>Please assess how many application fields you see and how large these fields are. Name a maximum of twenty application fields specifically and choose those you consider most relevant (if more, please only number of fields).</i>	Mit sehr geringen latenzzeiten z.b. bei firmen mit BI-abteilungen, die nicht bereit sind auf "heiße datensätze" zu warten (auch wenn es nur 10-20s sind) → in seiner welt werden die daten aus dem Hadoop system in exasol/hana zu laden andere firma für online-advertisement; wollen werbung zeigen und wirklich schnell das profil des kunden der auf der homepage ist klicken hauptsächlich dashboards, BI, online-marketing insgesamt wird sich big-data weiter ausbreiten, glaubt er auch; wird eher explosionsartig mehr als weniger
5	What are potential customer groups for in-memory technology, and specifically in-memory databases? <i>Please explain briefly why you consider this customer group relevant</i>	
6	Please make an educated guess on the intensity of research and development and the type of development. Können Sie abschätzen wie intensive Forschung & Entwicklung für relationale IMDB betrieben wird? Die Intensivitätsgrade sind: - mittel (Grundlagenforschung) - hoch (Anwendungsfelder) - niedrig (Kostenreduktion) - sehr niedrig (reife Technologie) Scale: - Medium (basics) - High (applications) - Low (cost reduction) - Very low	Forschung eher nicht, aber entwicklung in jedem fall: gerade bei HANA, die probieren phora (neueste strategie) an Hadoop anzubinden, da ist immense entwicklungstechnologie; exasol genauso IMDB die open source sind auch sehr Forschung eher nicht Entwicklung ja v.a. in hinsicht auf anbindung auf hadoop (da hinkte SAP hinterher)

	<i>Please explain your answer briefly</i>	
7	Please assess how long you think it will take for the technology to reach the stage of maturity, or whether it may already be mature. Wie lange wird es Ihrer Meinung nach brauchen bis die Technologie reif ist (d.h. weniger Forschung nötig, Anwendungsfelder bekannt etc.): Scale: - Long term (7-15 years until maturity) - Medium term (2-4 years until maturity) - Short term (>1 year until maturity) - Already mature <i>Please explain your answer briefly</i>	Manche sachen ist es ziemlich unreif (wie anbindung an andere technologien) In anderen bereichen ziemlich reif (queries) Reife am markt: troubleshooting, fehler, dokumentation → leider eher mittelreif Dauert schon noch 1-2 jahre bis das bei kleineren firmen überhaupt als option ankommt
8	Please assess in how far in-memory technology, specifically in-memory databases are available for potential customers. If you think availability differs in certain regions (e.g. Europe, USA, Southeast-Asia, etc.), please specify. Inwiefern sind IMD schon für potenzielle Kunden verfügbar? - Sehr eingeschränkt verfügbar (e.g. nur ausgewählte Test-Kunden) - Es gibt Wartezeiten für potenzielle Kunden (z.b. invite-only beta) - Marktorienterte Verfügbarkeit (gezielt für best. Kunden verfügbar) - Massenproduktion (für jeden verfügbar) Gibt es regionale Unterschiede? (Europa VS USA vs Asien etc.) Scale: - Very limited availability - Waiting periods for potential customers - Availability is market oriented - Mass production	
9	What are your thoughts on the business potential databases using in-memory technology?	Viele kunden sagen: wenn es was geben würde würden wir das auch nehmen

	Wie schätzen Sie das Geschäftspotenzial von IMDB ein?	HP hortiga? Auch ein kunde Sind alle zu teuer, sobald sie günstiger sind gibt es da potenziale, dann müssen sie ihr geschäftesmodell ändern
10	Please describe the influencing factors and potential drivers you consider relevant for the future development of databases using in-memory technology. Welche Faktoren könnten die zukünftige Entwicklung von IMDB beeinflussen? Relevant könnten sein: i) Marktentwicklungen (auch: Konkurrenz!) j) Verbreitung der Technologie k) Technologische Komponenten und Aspekte der Technologie l) Etc. Relevant for the future development of this technology are factors and drivers concerning a) market development b) diffusion of the technology c) technological aspects d) and any other potential aspects you may think of <i>Please feel free to name several influencing factors and drivers, even if you are not entirely sure.</i>	- Price - Anbindung an Hadoop - Documentation - Transparenz am markt (man kann nicht einfach hin gehen und sagen ich will hana kaufen, und dann ist es in ner halben stunde abgeschlossen) - Sogar weniger technologische aspekte - Hardware ist gut, es gibt bestimmt auch immer mehr hybrid-DB geben (hat HANA mit smart-data access schon!)

Interview Notes E9

Not available, please contact the researcher for more information

Attachment B: Interview Transcripts Phase R2

Before reading the transcripts, please note:

The transcripts have been cleaned up to exclude any filling sounds and words such as “Ähm”, “Mhm”, “OK”, “Aha”, “Ja”, “Ach”, “Also”, “Gut”, “In Ordnung”, “Genau” – unless relevant for understanding the content and context of the conversation.

Moreover, unless relevant for content and context, laughter and other conversational sounds are omitted.

Colloquial abbreviations such as “’ner” for “einer” were corrected into their grammatically correct versions.

Conversational welcoming remarks and introductions, as well as thank-yous and goodbyes are not included in the transcript, as they were not deemed relevant content-wise to the research. Such summarised sections or sections with sounds are marked in [blue], while inaudible sections from the audio tapes are marked in [yellow].

The numbers in brackets demonstrate the duration of the spoken paragraph in minutes and seconds: [mm:ss].

Interview Guideline Phase R2

QUESTIONS / REPLIES
QB1: Nachfrage von Echtzeit-Reaktion bzw. Echtzeit-Analyse und Beschleunigung von Services
Q1a: Wer braucht Ihrer Erfahrung und Einschätzung nach überhaupt eine Echtzeit-Analyse von Daten?
Q1b: Glauben Sie, dass Sie in 10 Jahren noch die gleiche Aussage treffen würden? Was hat sich bis dahin vielleicht verändert? Warum ja/warum nein?
Q1b: Beobachten Sie Trends, welche Arten von Unternehmen (oder vielleicht sogar auch Regierungen?) neue Systeme zur Echtzeit-Analyse bei sich einrichten und anwenden möchten? Hat sich hier in letzter Zeit etwas verändert?
Q1c: In welchen Bereichen gibt es noch nicht ausgeschöpfte Potenziale für Echtzeit-Analysen? Welche Bereiche könnten innerhalb der nächsten 10 Jahre an Relevanz gewinnen?
Q1d: Wie schätzen Sie die weitere Entwicklung der Nachfrage nach Echtzeit-Analysen und der Beschleunigung von Services innerhalb der nächsten 10 Jahre ein?
Q1e: Könnten Sie sich eine Situation vorstellen, weswegen die Nachfrage nach derartigen Systemen (wie relationale IMDB) innerhalb der nächsten zehn Jahre stagniert oder gar wieder sinkt?
QB2: Kosten relationaler in-memory Datenbank-Systeme (Kunden)
Q2a: Aktuell sind die meisten relationalen IMDB ja lizenzierte Systeme. Dadurch sind diese aber für viele Kunden recht teuer. Können Sie sich ein anderes Geschäftsmodell für relationale IMDB vorstellen, das nicht auf Lizenzgebühren basiert?
Q2b: Können Sie sich ansonsten Umstände vorstellen, weshalb relationale IMDB für die Kunden günstiger werden?
Q2c: Und andersherum: Warum könnte es passieren, dass relationale IMDB für den Kunden teurer werden?
Q2d: Wie schätzen Sie die zukünftige Kostenentwicklung für relationale IMDB innerhalb der nächsten zehn Jahre insgesamt ein? Sehen Sie mehrere verschiedene Optionen?
QB3: Entwicklung von Open-Source Lösungen für relationale In-Memory Datenbanken
Q3a: Wer hätte ein Interesse daran, eine Open-Source Lösung für relationale IMDB zu entwickeln? Können Sie sich vorstellen, dass das demnächst passiert, oder kennen Sie gar erste Beispiele?
Q3b: Können Sie sich vorstellen, dass die großen Anbieter relationaler IMDB (wie SAP mit Hana) selbst anfangen Open-Source Lösungen zu vermarkten? Warum ja/nein? Wie könnte das aussehen?
Q3c: Könnte es sein, dass in zehn Jahren sowohl Open-Source als auch lizenzierte relationale IMDB nebeneinander existieren? Warum könnte sich so ein Modell ergeben? Welche Kundengruppen und Geschäftsbereiche fänden welches Variante interessant?
Q3d: Nehmen wir nun einmal an, innerhalb der nächsten zehn Jahre hätten verschiedene Open-Source Lösungen die herkömmlichen lizenzierten relationalen IMDB vom Markt verdrängt. Wie würden Sie sich die Entwicklung zu diesem Endstadium vorstellen?
Q3e: Unter Betrachtung von allem, was wir bisher besprochen haben: Wie stellen Sie sich die Entwicklung von Open-Source Lösungen für relationalen IMDB innerhalb der nächsten 10 Jahre vor?
QB4: Gesetzliche Regulierung (national/EU/international) bezüglich Schutz der Privatsphäre, Daten-Sicherheit und Standardisierung/Normierung
Q4a: Die EU hat ja kürzlich eine neue Regelung zum Schutz von persönlichen Daten und der Privatsphäre vorgestellt, die bald in Kraft tritt. Ganz allgemein: Was kommt hier auf Unternehmen zu, die Daten sammeln, speichern und analysieren? <i>(Bitte um kurze Zusammenfassung)</i>
Q4b: In Verbindung mit Nutzerdaten von Services ist der Zusammenhang zwischen diesen gesetzlichen Regulierungen und möglichen Einschränkungen beim Sammeln/Speichern/Analysieren von personenbezogenen Daten relativ gut nachvollziehbar. Welche möglichen Einschränkungen sehen Sie Firmen-intern, beispielsweise bei Produktionsprozessen & Inventar?
Q4c: Das Schlagwort „Industrie 4.0“ umfasst grob gesagt eine Effizienzsteigerung, Digitalisierung und weitere Automatisierung von Geschäftsprozessen durch Big Data. Welche Möglichkeiten können Sie sich von gesetzlicher Seite vorstellen, die diese Automatisierung stoppt? Warum würde die Regierung diesen Prozess fördern?
Q4d: Welche Maßnahmen von Regierungsseite können Sie sich sonst noch vorstellen, um die Potenziale von Big Data optimal zu nutzen?

Q4b: Fallen Ihnen sonst noch Regierungsmaßnahmen und Gesetzesvorschriften ein, die die Entwicklung von Big Data Analyse im Allgemeinen innerhalb der nächsten 10 Jahre behindern ODER begünstigen würde?
QB5: Weiterentwicklung des Verbindungsnetzwerks <i>Intro: Im Internet of Things sollen über Sensoren ja ständig eine Vielzahl an Daten übertragen werden. Dies wird die Relevanz von Big Data Analyse noch einmal in ganz andere Dimensionen bringen. Aber das Netzwerk muss für die Übertragung derartiger Datenmengen bereit sein.</i>
Q5a: Welche Technologien könnten in 10 Jahren diese Last an Daten stemmen?
Ideen: - Flächendeckendes WiFi (e.g. Google's Project Loon) - 5G Netzwerk - Bluetooth Beacons - Infrarot-Schnittstellen - Neue Technologie: LiFi (über Licht; bis zu 2400x schneller als WiFi) - Kombination aus den 4
Q5b: Inwiefern könnten einzelne Unternehmen interne Lösungen für den Datentransfer finden? („Insel-Lösungen“)
Wie sähe es dann bei Konsumenten-Produkten aus? (z.B. Smart Homes; Wearables)
Q5c: Können Sie sich ein Szenario vorstellen, in dem sich das Internet of Things zu schnell ausbreitet, das Netzwerk in der Entwicklung aber NICHT mithalten kann? Wie sähe das aus?
Was wären die Konsequenzen für Unternehmen im Bereich Big Data Analyse?
Q5d: Woher könnten Impulse für eine Verbesserung des Netzwerks kommen?
QB6: Anzahl der Sensoren im Bereich IoT
Q6a: Was halten Sie von den Prognosen zum Anstieg der Sensoren und der explosionsartigen Verbreitung von „Big Data“ und dem IoT? Ist das ganze nur ein Hype?
Q6b: Welche Entwicklungen könnten die Verbreitung von Sensoren stagnieren lassen? (Gesetze, Rohstoffe, Naturkatastrophen?)
Q6c: Warum könnte Big Data weiter wachsen, wenn die Prognosen zur explosionsartigen Steigerung der Sensoren innerhalb der nächsten zehn Jahre nicht stimmen?
Q6d: In welchen Bereichen sehen Sie die größte Relevanz von Sensoren und IoT Technologie?
Q6e: Welche Überraschungen könnten Sie sich für die Verbreitung des Internet of Things und Big Data Analyse innerhalb der nächsten zehn Jahre vorstellen?

Transcript E4

Date: 29.01.2016

Duration: 53:31 [mm:ss]

Modus: face-to-face interview

Interviewee:

Expert 4 [E4]: Jörg Frank; Consultant and Director of Systems Development in international consulting firm Syncwork in Berlin

Interviewers:

Julia Reichelt, Researcher [R]

Victoria Götz, Thesis Supervisor [TS]

[00:00-02:31]: *[Conversational introductory remarks]*

R [02:31-02:40]: Vielleicht können Sie erstmal kurz zusammenfassen, wie Sie sich mit relationalen in-memory Datenbanken [RIMDB] beschäftigen?

E4 [02:40-03:55]: Grundsätzlich von meinem Wirkungsspektrum: Wir sind hier ein Beratungshaus, wo wir auch unter anderem die Kunden im Bereich Warehousing und Analytik beraten, Projekte umsetzen, also beraten. Ganz oft die Projekte konzipieren, umsetzen und die Systeme zu durchlaufen. Da ist natürlich Analytik nachher immer ein Thema, genauso wie Performance. Und typischerweise klassisch sind das relationale Datenbanken, die dann heute ergänzt werden um RIMDB. Also ich habe jetzt vor kurzem für ein BI-Projekt, das konkret mit SAP HANA gemacht wurde, und das andere ist die ganze Welt der Hadoop, NoSQL, die auch noch heute dazu ist, das ist nochmal eine andere Richtung. Und dann die Verschmelzung all dieser Technologien heute. Ich glaube, dass sozusagen diese Grenzen die man heute so hat fast verschwinden werden und nur noch schwer zu fassen sind.

R [03:55-04:00]: Das ist natürlich auch sehr interessant für uns.

E4 [04:01-04:03]: Also, das ist das was ich mache.

R [04:04-04:10]: Also Sie machen also beides, also in-memory und relationale Seite, als auch NoSQL?

E4 [04:10-04:22]: Ja genau. Wobei der Schwerpunkt ist schon auf relationalen Datenbanken, weil das ist das was in der Welt noch am weitesten verbreitet ist.

R [04:23-04:55]: Dann werden wir gleich schon konkreter für den ersten Fragenblock. Da geht es jetzt erstmal ganz allgemein um die Nachfrage von Echtzeit-Analysen und Echtzeit-Reaktion, und auch wenn man jetzt Services speziell anschaut, dass man die immer mehr beschleunigen will. Zur ersten Frage, auch aus Ihren eigenen Projekterfahrungen – wer braucht so etwas überhaupt, wer sind die wichtigsten Kundengruppen?

E4 [04:55-6:51]: Ich glaube das geht über alle Kundengruppen. Ich bin in relativ vielen Industrien tätig, sowohl Healthcare als auch Telco bis hin zu Bergbau, und in den Industrien gibt es überall Ansatzpunkte wo Echtzeitanalysen interessant sind. Das eine ist in... Oder meistens sind das Prozessabläufe: Analytik für Prozessabläufe. Nicht nur Prozesse von Maschinen, sondern auch Arbeitsabläufe, z.B. aus der Telekommunikation wo wir das seit Jahr diskutieren, wo man sowas gebrauchen kann, ist der Ablauf von: Sie gehen in den Laden, und kriegen einen neuen Handyvertrag, dann ist die... dann kriegen Sie eine SIM-Karte und ein Telefon, dann wollen Sie relativ schnell telefonieren. Da sind so zwischen dreißig bis fünfzig Prozessschritte in einem Telekommunikationsunternehmen notwendig für das fulfillment bis Sie sozusagen das benutzen können. Und wenn man sowas monitoren will und relativ schnell eingreifen möchte in so einen Prozess, dann bin ich bei einer Echtzeitanalyse, Maschinenanalysen. Und das ist glaube ich über die Industrien querbeet gibt es Anwendungsszenarien, aber es gibt aber auch weiterhin ganz viele, oder die Majorität der Anwendungsszenarien braucht es eigentlich nicht. Es gibt ausgewählte, ganz viele Analysen, wenn ich mein Geschäft analysiere und sowas, dann sind Echtzeitanalysen zum sogar Gift, weil ich mindestens einen Tag eine stabile Analytik brauche. Sie haben das Problem – ich mache ne Analyse jetzt, dann gehe ich zu einem Kollegen hin, ne Stunde will ich mit dem diskutieren, der schaut sich die an, die Analyse ist anders, weil Sie in Echtzeit alles haben, dann haben Sie ein Problem auf einmal. Es ist nicht nur alles gut, sondern das ist wirklich sehr anwendungsbezogen.

R [06:51-07:06]: Und glauben Sie, dass sich dieses... das verändert, also gerade wenn Sie jetzt sagen der Großteil braucht es eigentlich nicht? Wird es da einen Änderungsschritt geben, vor allem jetzt mit dem Zeithorizont zehn Jahre den wir aufspannen wollen?

E4 [07:06-7:33]: Ja glaube ich schon. Und zwar vor allen Dingen weil die klassische Analytik, die tagesgenaue sagen wir einfach mal, die beherrscht man inzwischen relativ gut und ist auch schon weit verbreitet. Jetzt ist eigentlich der nächste Schritt neue analytische Felder zu erschließen. Und da ist die Echtzeitanalytik ein Teil davon. Das genauer zu erschließen.

R [07:33-07:35]: Und was gäbe es noch für Bereiche?

E4 [07:35-08:11]: Prognosen. Vorhersagen. All das ist ja auch ein großer Trend, dass ich jetzt nicht nur retrospektiv schaue, sondern möglichst gute Prognosen mache, und das ist halt eben auch etwas was... Prognosen brauchen weniger Echtzeit, also je nachdem. Muss man etwas vorsichtig sein, brauche ich etwas weniger. Aber gerade was jetzt so Prozesse zum monitoren und aktiv einzugreifen sehr kurzfristig, dafür brauche ich Echtzeit, und da wagt man sich jetzt einfach ran weil man den Rest beherrscht.

R [08:11-08:25]: Und sehen Sie da auch schon Trends, bei den Unternehmen selber, dass die da wirklich massiv investieren und versuchen dieses Thema einzubauen, oder ist das eher ein schleppender Prozess?

E4 [08:26-08:52]: Also da wo ich unterwegs bin ist es ein Stück schleppender Prozess würde ich sagen, da wird jetzt nicht so mit Hochdruck... Das ist natürlich, das zu etablieren ist nicht einfach, also [inaudible]. Im Endeffekt.. und der Businesswert der dagegen steht, den muss ich halt irgendwie rechnen. Und da sind die business cases heute zum Teil noch schwierig, [inaudible].

R [08:54-09:03]: Und würden Sie da auch sagen das wird sich aber verändern, weil es zum Beispiel günstiger wird, oder weil es einfach der Standard wird zum Beispiel?

E4 [09:04-10:13]: Ich denke, also der große Teil der Kosten basiert ja nicht auf hardware oder software Kosten sondern die Projektkosten, die sehr umfangreich sind. Und da hat man noch nicht so den großen Erfahrungsschatz. Und der wird sich aufbauen, dann wird es besser kalkulierbar und günstiger werden. Und da wird dann einfach sozusagen dort, der wird da noch um den best [inaudible] früher zu sehen sein, dann macht man's eher. Das andere ist dann noch sozusagen wenn die Konkurrenz es macht, dann entsteht dann so ein Druck: Man hat einen Schaden durch unterlassen. Zum Beispiel, Vodafone führt ein Prozessmonitoring ein, und bei Vodafone spricht sich rum, dass wenn ich dort die SIM-Karte krieg, dann kann ich immer 30 Sekunden später das erste Telefonat führen – Telekom macht das nicht, und es spricht bei denen dauert es manchmal fünf Minuten! Dann entsteht ein image-Schaden und allein dadurch kommt ein business case zustande, dass ich es machen muss. Also der ansonsten nicht ist, das wird dann so ein Marktdruck der da entsteht.

R [10:14-10:19]: Und glaube Sie, der Marktdruck wird stark steigen innerhalb zehn Jahren, oder braucht das noch länger?

E4 [10:19-10:39]: Zehn Jahre, ist schon lang. Ich glaube wenn wir jetzt mal zehn Jahre zurück denken, was da war und was heute ist, dann sind die Unterschiede schon massiv. Zurzeit beschleunigt sich ja eher die Evolution im IT-Bereich. Da wird das schon lang sein.

R [10:40-11:00]: Aus dem Vorherigen würde ich jetzt schließen Sie sehen da schon noch große Potenziale für die nächsten zehn Jahre, also dass da schon noch sehr viel zu machen ist was Echtzeit-Analysen und auch – Prognosen betrifft.

E4 [11:00-11:04]: Ja, also ich glaube das kommt dann immer mehr.

R [11:05-11:26]: Und konkret zur Nachfrage, also dem was wir als Driver festgelegt haben, würden Sie auch sagen, dass sich das wirklich... das das auf jeden Fall steigt, oder könnten Sie sich auch irgendwas vorstellen, wonach die Nachfrage stagniert oder sogar wieder sinkt.

E4 [11:26-11:28]: Für Echtzeit-Analysen jetzt genau? Also die Analytik [inaudible]?

R [11:28-11:33]: Genau, und diese Schnelligkeit..

E4 [11:34-11:54]: Ne, also unsere Gesellschaft beschleunigt, und damit wird das auch, das wird sich da stark wieder spiegeln. Also ich glaub, dass das sozusagen von diesen Anforderungen kann ich mir heute nicht vorstellen, dass es sozusagen zu einer Entschleunigung in den Prozessen kommt.

R [11:54-12:01]: Aber unendlich weiter führen kann man das natürlich nicht, ne?

E4 [12:01-12:29]: Ne, das muss ich ehrlich sagen, das weiß ich gar nicht genau. Vor hundert Jahren hätte sich ja auch keiner vorstellen können was wir heute alles so machen. Kommunizieren mit der ganzen Welt in Echtzeit, was hatten wir dann, die Telegraphie mal erfunden, und nur Telefonieren, und deswegen... Soweit kann ich nicht denken.

R [12:29-12:52]: Und können Sie sich vielleicht vorstellen, dass es dafür manche Bereiche gibt, die praktisch dann massive Vorreiter sind, und andere dadurch zurück fallen? Also Sie sagen auch gerade wo Sie unterwegs sind kommt es Ihnen doch eher schleppend vor, gibt es vielleicht manche die da viel viel schneller sind?

E4 [12:54-14:10]: Eigentlich die klassischen Industrien könnten sich vielleicht ein Stück zurück halten an der Stelle. Sicherlich sind die, ich sag mal, Internetunternehmen, also Google etc., werden versuchen dort auch weit vorne zu sein, weil sie einfach technologisch weiter sind, und ich glaube nicht nur technologisch, sondern sagen wir auch mal vom Mindset her so weit vorne sind. Ich denke, dass die Industrie sozusagen ganz stark in automatisierte Fertigung, Entwicklung geht, was auch ganz viel Realtime-Analytik braucht. Nur die Frage ist, wenn man jetzt mal ganz weit denkt, inwieweit der Mensch nachher die Analytik noch benutzt. Halte ich nicht für ausgeschlossen, dass irgendwann Maschinen die Analytik auswerten, und wieder mehr oder weniger autonom sind. Also sowas wie vielleicht nicht in zehn Jahren, aber irgendwann kann man sich vorstellen, dass es sowas wie autonome Fertigung gibt wo der Mensch eigentlich nix mehr zu suchen hat.

R [14:10-14:14]: ...also dann wirklich in Richtung künstliche Intelligenz für Industrie.

E4 [14:14-14:39]: Ja ja! Also sagen wir mal... Die Automobil-Industrie ist sicherlich relativ weit vorne was Robotik betrifft, und diese Produktion, wenn man das ein Stück weiter denkt, und die Maschinen das auch noch auswerten und das läuft und selber alles machen. Dann sind wir eigentlich etwas, was wir uns heute nicht vorstellen können.

R [14:39-14:42]: Die volle Automatisierung sozusagen.

E4 [14:42-15:11]: Ja! Also das... Ich habe mal einen interessanten Vortrag gehört, das selbst in so archaischen Industrien wie der Landwirtschaft ist es so, dass heute da gibt es selbstfahrende Maschinen. Also auf den Traktoren sitzt keiner mehr. Der fährt alleine aufs Feld, macht das alles, und selbst das könnte voll automatisiert sein.

R [15:12-15:24]: Wirklich spannend. Auch gerade im Bereich selbstfahrende Fahrzeuge glaube ich, dass Landwirtschaft da ein interessanter Bereich ist.

E4 [15:24-15:27]: Und das gibt's schon!

TS [15:27-15:32]: Und die Bauern sitzen nicht im Feld sondern im Haus und steuern das alles.

E4 [15:32-15:33]: Nein, das braucht der auch nicht mehr!

TS [15:33-15:34]: Überwachen...

E4 [15:34-16:17]: Ja, vielleicht schaut der abends ah, sie sind alle wieder da. Aber es ist wirklich so, so ein... Diese alte Industrie, wo man denkt das ist so ganz hinten dran, aber nein, selbst da gibt es diese Ansätze. Und das schaut ja dann auch wieder hin zur Analytik. Treibe ich das gerade in der Automatisierung... Im Endeffekt: Kriegen die Maschinen dann schon einen Wetterbericht, und überlegen selber wo sie jetzt hin fahren, weil in ner halben Stunde regnet es dort? Genau das ist Online-Analytik

R [16:18-16:57]: Das ist schon sehr spannend. Dann kommen wir erstmal noch zu einem anderen Bereich, und zwar den Kosten. Wir haben jetzt auch schon in vorherigen Interviews öfter mal gehört, die Kosten sind so hoch, da die meisten RIMDB mit Lizenzmodellen arbeiten. Erstens, würden Sie dem zustimmen? Haben Sie da auch

schon Erfahrungen? Und zweitens, können Sie sich andere Modelle vorstellen wie das System praktisch funktionieren könnte ohne diese Lizenzen? Also gibt es zum Beispiel andere Geschäftsmodelle?

E4 [16:57-21:08]: Das gibt's, ja. Also... Vielleicht sehe ich das auch ein bisschen einfacher. So eine RIMDB ist auch nur eine Datenbank, oder ein Stück software, was die Daten im Hauptspeicher hält, auf einer Festplatte. Und natürlich kann man sich vorstellen, es gibt ja freie, open-source, relationale Datenbanken. Die software ist ja nicht so krass unterschiedlich – es gibt ja auch open-source in-memory Datenbanken, und es kann sich natürlich so ne Community weiter entwickeln. Sodass man neben den kommerziellen Anbietern auch open-source, oder sozusagen freie Anbieter hat. Und wenn man sich die normale relationale Welt anschaut ist das ja schon so. Man hat Oracle, IBM, Microsoft als kommerzielle Anbieter dort, und man hat dann eine MySQL, oder was auch immer nebdan als open-source Datenbank, die frei verfügbar ist. Und das kann ich mir vorstellen, dass das im in-memory Bereiche auch so ne Co-Existence gibt. Weil es wird weiterhin Unternehmen geben, wenn das unternehmenskritische Datenbanken sind, die wollen einen Hersteller haben, den sie verklagen können, was auch immer. Dafür bezahlen sie. Und es gibt andere, die kostenlose Anbieter haben wollen. Ich denk... das sehe ich gar nicht.

Der Grund was natürlich heute immer noch ist, ist die hardware, auch halt eben für große Datenbank brauch ich entsprechend große Hardware, weil da natürlich die Preise auch... aktuell RAM wird günstiger, damit werden die Maschinen günstiger. Andererseits wird Datenvolumen immer größer. Ich mache ja schon lange Datawarehouse, irgendwann war ein Terabyte Datawarehouse richtig groß, und inzwischen ist es klein. Jetzt ist ein Petabyte ein wirklich großes Warehouse. Ein Petabyte RAM gibt es glaub ich momentan nicht, das kriegt man gar nicht zusammen. So gibt's also sozusagen: Die Datenmenge nimmt stetig zu, auch die die verarbeitet werden möchte, oder die man verarbeiten will, gleichzeitig sinken die Preise, aber durch die Zunahme frisst das sich gegenseitig wieder ein bisschen auf.

Ich hab ja angefangen und gesagt, ich glaube, dass die Sachen zum Teil gegeneinander sich irgendwann... Also die Grenzen werden nicht mehr klar sein, oder vielleicht ganz verschwinden. Also was ist jetzt an einer in-memory Datenbank anders als an einer anderen nicht-in-memory Datenbank. Könnten Sie sagen, das eine liegt im Hauptspeicher, das andere nicht. Aber ansonsten, im interface, das ist alles das gleiche. Warum soll ich jetzt überhaupt noch wissen was meine Datenbank dort macht? Es gibt seit längerem so hot-code data-approaches. Das heißt hot data ist welche, die ich schnell zur Verfügung haben muss, oder häufig, die auf sehr schnelle Medien packe, RAM ist einfach die schnellste, was wir haben aktuell. Vielleicht zu warm data brauche ich häufig, pack ich auf schnelle Platten, also SSD-Platten zum Beispiel. Und zu cold data: riesige Datenmengen, wo ich aber selten rein schaue, und wenn ich rein schaue, dann kann ich auch mal ein paar Stunden warten, das kann man auf große Platten packen.

In Zukunft wird man glaube ich von intelligenten Systemen erwarten, dass das System kapiert OK, das ist hot data, das muss ich in den RAM packen, dass ist cold data, das kann liegen bleiben. Dann hat man so hybride Systeme. Gibt's zum Teil auch schon! Die Hersteller sind genau daran am Arbeiten, um auch einfach diese riesigen Datenmengen zu verarbeiten.

R [21:09- 21:35]: Einen Moment bitte, zur Vorsicht notiere ich doch immer noch etwas mit falls mit dem Audio etwas nicht passt. So: Das ist auf jeden Fall sehr spannend; Hybridmodelle, das passt auch schon zu den nächsten Fragen: das wäre dann eine Alternative für Sie?

E4 [21:35-22:00]: Ja, das glaube ich. Ich glaube, das wird irgendwann... Das will ich nicht mehr wissen, wo das System das hin speichert. Ich habe Anforderungen, wie schnell muss ich diese Daten auswerten, verarbeiten können, und das System packt das da eigentlich dort hin, wo es in angemessener Geschwindigkeit verarbeitet werden kann.

R [22:00-22:23]: Und können Sie sich auch ein Szenario vorstellen – es geht ja jetzt um die Kosten – wo es für die Kunden teurer werden würde? Also einerseits natürlich die Datenmenge, aber prinzipiell gehen ja wahrscheinlich die Prozesse dahin, dass man versucht die Kosten zu senken.

E4 [22:24-23:25]: Naja, also die Entwicklungen der letzten Jahrzehnte haben ja gezeigt, dass hardware sukzessive günstiger wird. Und teilweise bleiben die Systeme gleich teuer weil sie einfach leistungsfähiger werden. Also ich rechne nicht damit dass in den nächsten Jahrzehnten, dass hardware-Preise auf einmal steigen. Ich glaube eher das wird sich weiter so bewegen. Was natürlich ist, ist dass die Datenmenge steigt, und dann müssen die Systeme größer werden, deswegen könnten die Kosten auch ein Stückweit stabil bleiben; sozusagen eher Verfall des hardware-Preises wird durch die höheren Anforderungen an Datenmengen und auch die Verarbeitung der größeren Datenmengen – es ist ja nicht nur dieses sie speichern es, sie müssen ja immer genügend Speicher am

Ende [inaudible], ich muss ja auch wieder prozessieren. [inaudible] ...dass ich sagen würde, OK das Gigabyte RAM, das in zehn Jahren teurer wird glaube ich nicht.

R [23:26-24:04]: Wenn Sie sich jetzt aber mal überlegen wie diese ganzen Prognosen, zum Beispiel von Gartner, wie viele Trillionen an Internet-of-Things Devices es geben wird, und wie viele Daten dann da entstehen, wenn das wirklich so explosionsartig passieren wird, würden dann die Systeme da noch mithalten, und würde das sich praktisch – könnte es sein, dass da auch dadurch dann die Kosten extrem steigen weil man sagt OK, wir wollen zwar alle diese Devices haben, aber unsere Technologie ist aber leider noch nicht so günstig und so verfügbar.

E4 [24:04-25:54]: Das eine sind ja die Devices an und für sich. Also, ich vergesse immer die Zahl so ein bisschen, ich glaub in einem modernen Auto hab ich inzwischen 80 Datenbanken drin, irgend sowas in der Größenordnung, so embedded systems, die heute noch relativ – die stecken im Auto. Die kommen da nicht so häufig raus, die können ausgelesen werden wenn ich in ne Werkstatt fahre, aber werden nicht kommuniziert. Und das ist schon ne Menge. Und ich weiß auch gar nicht wie viele Gigabyte Daten da pro Auto gespeichert werden, das sind aber wirklich viele. Die werden aber alle nicht zur Analyse gebracht. Und das ist diese wahnsinnige Datenmenge, die jetzt im Internet-of-Things [IoT] entsteht erstmal; also wie werden dann die Datenflüsse sein, wie viel davon wird verarbeitet, wie viel wird als Rohdaten verarbeitet, wie viel wird im Endeffekt erstmal analysiert und in einer komprimierten Form, in einer aggregierten Form verarbeitet. Deswegen ist es... kann man das nicht so platt sagen es wird so einen Datenberg werden und wir müssen den komplett verarbeiten, und dann geht gar nichts mehr, das überfordert uns total. Sondern, das ist das was wir heute im Big Data Umfeld schon machen – wenn ich heute ein Hadoop-System an soziale Medien anschließe oder sowas, dann prozessiere ich zwar die ganzen messages, die da kommen – üblicherweise weitergeben tu ich aber schon exakte Extraktionen, und denke das wird auch der Weg sein, den man zumindest zuerst mal suchen wird, um die Masse erstmal kleiner zu machen und auch beherrschbar zu kriegen.

R [25:55-26:04]: Also dieser Verkleinerung der Datenmenge ist aber auf jeden Fall, würden Sie schon als sehr wichtig ansehen, damit man da nicht so einen riesen Datenwust hat?

E4 [26:04-26:57]: Ja! Das ist ja auch, also es gibt ja auch einmal sozusagen der reine Datenwust, den ich dort habe, das ist einfach ein Mengenproblem, auch zu prozessieren. Ich spreche auch gern von intellektuellen Problemen sie zu verstehen. Also je detaillierter, je umfangreicher, gerade Messdaten sind – Messdaten haben Messfehler drin, also sind nicht sauber – was machen sie mit dem Einzeldatensatz eigentlich. Eigentlich können Sie sie nur gut betrachten wenn sie einen Durchschnittswert, oder was auch immer, also so eine aggregierte Form, da können Sie sie gut weiterverarbeiten. Das entlastet auch intellektuell – weil vor der Herausforderung stehen wir auch. Weil diese Masse an Daten, und Datenzusammenhängen nachher intellektuell zu beherrschen ist meiner Meinung nach genauso schwierig wie sie mit den Maschinen zu beherrschen.

R [26:58-27:07]: *[brief pause for writing notes]*

R [27:07-27:19]: Also Sie sehen das praktisch mehr al sein Problem des Datenmanagements sage ich jetzt mal als der konkreten Kostenfrage.

E4 [27:19-27:23]: Ja, das glaube ich schon.

R [27:23-27:49]: Dann können wir auch gleich schon zum nächsten Block gehen. Da haben wir ja eigentlich auch schon drüber geredet, und zwar diese Open-Source Solutions. Kennen Sie da schon Beispiele? Also ist das so, da bin ich mir jetzt nicht sicher, es gibt relationale Datenbanken und in-memory Datenbanken open-source, aber nicht als eins?

E4 [27:49-28:47]: Oh Gott, wie ist denn das? MySQL hat glaube ich inzwischen auch schon ein Stück in-memory Komponente, bin mir aber nicht hundert Prozent sicher. Es gibt so ein paar, H2 zum Beispiel, die in-memory open-source Datenbanken sind. Es gibt bestimmt eine Handvoll von denen, die sind teilweise einfach im universitären Umfeld entstanden, und teilweise natürlich in Projekten, in Unternehmen, die dann einfach daraus open-source gemacht haben, wo sie sich ja erhoffen, dass es von der Gemeinde weiterentwickelt wird. Das ist ja immer das: Warum mache ich was open-source? Es gibt da so ein paar Gründe, aber ein wichtiger Grund ist, ich hab für mich da was gemacht, und möchte das weiter entwickeln, und wenn ich es der Gemeinschaft zur Verfügung stelle arbeitet die Gemeinschaft dran, und ich profitiere davon.

R [28:48-28:53]: Also Sie können sich da schon auch vorstellen, dass größere Unternehmen in die open-source Richtung gehen würden?

E4 [28:53-29:05]: Ja, die großen Internetunternehmen haben es uns vor gemacht, also wieder Google etc., große Teile von ihren Softwareentwicklungen als open-source dann zur Verfügung stellen.

R [29:05-29:18]: Und auch jetzt beim konkreten Beispiel SAP, oder auch Oracle, IBM, können Sie sich da vorstellen, dass da auch mit eine Tendenz kommt, oder wollen die ihre Lizenzen behalten?

E4 [29:18-31:39]: Genau, wovon leben sie denn, da muss man das Geschäftsmodell der Unternehmen sehen, SAP hat schon mal eine Datenbank open-source gestellt, die SAP-DB, das relationale System das inzwischen über MaxDB zu MySQL gewandert, war ja dann eine open-source Datenbanken, oder eine frei verfügbare Datenbank, die SAP hergestellt hat.

Das hängt aber damit zusammen, dass es ja nicht Kern-Geschäftsmodell der SAP war, eine Datenbank zu vertreiben, sondern das war einfach eine alternative Datenbank aus eigenem Hause für ihr ERP System. Vielleicht um Oracle und IBM mal unter Druck setzten zu können um Lizenzierung, wie auch immer; gab es sicherlich etliche Gründe das zu tun. Und ich denke genau da liegt der wesentliche Punkt: Wenn die SAP, heute empfinde ich es so, eine SAP HANA als wirklichen Kern von ihrem Geschäftsmodell sieht, dann wird sie sie nicht frei machen. Ist SAP HANA nur ein Vehikel als Basistechnologie um ein ERP-System oder ein DB-System zu verkaufen, dann kann das durchaus bei der SAP so werden. Das sehe ich aber heute nicht so. Vielleicht kann das aber auch mal sein.

Eine Oracle, wenn wir die nehmen, ja Hauptgeschäft ist eine Datenbank zu verkaufen, wenn sie das nicht mehr tun, dann sind sie auf einmal arm. Dann wären sie schlecht beraten sozusagen in ein Lizenzmodell rein zu gehen, dass das Unternehmen nicht weiter existiert. Andere Unternehmen – wie gesagt, das ist ja das andere Geschäftsmodell, aber das sehe ich das bei den Datenbankherstellern heute nicht, ist ja sozusagen etwas frei zur Verfügung zu stellen, und rein an den Services das Geld zu verdienen. Das ist ja im Hadoop Umfeld, oder auch im Unix, Linux Umfeld wo es die freien Distributionen gibt, wo dann die Unternehmen entsprechende Services anbieten um die zu bündeln, add-ons zu liefern, und... ja.

R [31:40-32:06]: Und können Sie sich vorstellen, dass es zum Beispiel wenn sich jetzt ganz viele open-source Lösungen und drum herum Service-Anbieter für diese open-source Datenbanken ergeben, dass die entweder die zum Beispiel von SAP-HANA und diese von den großen Softwareanbietern, vom Markt verdrängen, oder die gezwungen werden sich...

E4 [32:05-33:03]: ...kann passieren! Also, wenn man Web-Server, Application-Server im Java-Umfeld, das ist jetzt wie lange her, fünfzehn Jahre her so rund, da gab es die lizenzpflichtigen Modelle auch von IBM und diversen anderen, Firmen, die jetzt gar keine große Rolle mehr spielen. Und dann kam so ein Druck aus der open-source Ecke, dass der Quasi-Standard auch im open-source Bereich liegt. Wenn man Unix sich anschaut heute ist der Hauptteil freie Unix-Distribution. Vor 20, 25 Jahren war der Hauptteil käufliche Unix Distribution. Also das kann passieren. Aktuell sehe ich das noch nicht so, aber wenn wir jetzt ein bisschen weiter raus denken, vielleicht in fünf Jahren – ja gute Frage.

R [33:04-33:06]: Könnte auch auf jeden Fall eine Option sein.

E4 [33:07-33:46]: Ja, wenn man jetzt zurück blickt, wie ist das an anderen Ecken passiert, Web-Applikationsserver ist ein Beispiel, Unix-Betriebssystem ist ein Beispiel, wenn man Parallelen dazu zieht, kann das da auch passieren. Die Datenbankwelt ist bisher ein bisschen davon verschont geblieben. Es gibt schon lange frei verfügbare Datenbanken, die aber die käuflichen nicht, zumindest in den großen Unternehmen, nicht verdrängt haben. Also in einem gewissen Marktsegment sind die gut dabei, aber die großen im Markt, [inaudible], IBM, eine Oracle.

R [33:47-34:51]: Dann hätten wir den Bereich auch schon abgedeckt. Dann kommen wir nochmal zu einem ganz anderen Bereich, und zwar gesetzliche Regulierung. Vor allem Privatsphäre, Sicherheit, und Standardisierung. Und ganz allgemein sind da unsere Gedanken: Wo kann überhaupt die Regierung eingreifen, was könnte passieren? Es gibt ja jetzt natürlich auch schon diese EU-Richtlinie zum Beispiel, und welche Bereiche sind da dann am meisten von Betroffenen, dass sie zum Beispiel nicht mehr so frei die Daten sammeln können?

Also wenn man jetzt mal Consumer-Applications anschaut ist es natürlich direkt nachvollziehbar, personenbezogene Daten, die müssen irgendwie geschützt werden, aber es gibt ja auch große Potenziale im Bereich Industrie, wie Sie ja schon meinten bei Autos zum Beispiel. Könnten die auch tangiert werden?

E4 [34:51-37:46]: Ja, also sicherlich, ich sage immer: ich glaube, dass wir wenn wir mal Autos anschauen, das kennt man ja gut, jeder hat da ein Stück weit mit zu tun, dann sind wir nicht weit davon weg, in zehn Jahren werden die Autos wahrscheinlich alleine durch die Gegend fahren, oder zumindest ein Teil, und die Autos werden auch alleine sich melden und sagen: Ich muss jetzt mal in die Werkstatt. Und dann, das impliziert aber natürlich, dass das grundsätzlich, also das Auto weiß wo es ist, also ist sicherlich auch irgendwo im Internet, wenn wir es dann überhaupt noch Internet nennen, aber sagen wir jetzt erstmal Internet. Das impliziert aber natürlich, dass irgendjemand wo ich, oder mein Fahrzeug zumindest ist, und ja, und was mache ich mit so einer – das ist sicherlich eine schützenswerte Information. Ich will ja nicht immer, dass die ganze Welt genau weiß wo ich bin, es sei denn ich will es der Welt mitteilen. Es gibt ja auch Menschen, die wollen allen immer mitteilen wo man gerade ist. Aber ich möchte das vielleicht auch mal nicht. Und das ist ein Punkt. Das andere ist natürlich... Das eine ist das Auto, die Vernetzung von all den Dingen. Genauso werden in zehn Jahren wahrscheinlich jegliches Device was wir haben, also sei es so einen Laptop, falls es den noch gibt, Tablet, Handy, Uhr – alles wird GPS haben, und alles wird eine Datenverbindung haben. Das heißt man kann da sozusagen aus der Masse der Devices, dies ja nun gebe, ganz viel raus ziehen über die Einzelperson. Und man wird es auch versuchen, das glaube ich. Und das ist ja von vielen Firmen, oder von Facebook oder so, ist es ja das Geschäftsmodell: ein exaktes Profilbild von einem Menschen zu erstellen, oder User, wie man ihn auch nennt, und mit diesem Profil, und dieses Profilbild dann zu verkaufen. Und gerade im Bereich Werbung, [inaudible], und gibt es ja schon viele aktuelle Einsätze, ich laufe an einem Laden vorbei und krieg dann die passende Werbung aus diesem Laden auf mein Handy. Oder ich steh schon vor dem Regal, und dann krieg ich die Meldung bitte rechts oben greifen, weil du weißt es zwar noch nicht, aber das willst du eigentlich jetzt kaufen. Das hört sich jetzt lustig an, aber dann kommen wir genau dahin: Wie viel von meiner Privatsphäre ist schützenswert, und wie viel möchte ich eigentlich geschützt haben. Das ist irgendwie so eine... ganz ganz schwierige Sache. Und da ist der Gesetzgeber gefragt. Weil die Unternehmen werden so viel nehmen wie sie können – ganz klar. Und der Gesetzgeber hat die Aufgabe sozusagen das Individuum zu schützen wo es sinnvoll ist.

R [37:47-38:07]: Und glauben Sie, dass das auch wirklich umsetzbar ist – also das Ideal wäre ja, dass die Gesetzgeber, gerade jetzt hier in Deutschland, möglichst viel der Privatsphäre schützen. Aber man sieht ja jetzt auch heutzutage schon, Facebook sammelt trotzdem unfassbar viele Daten, whatsapp, und so weiter.

E4 [38:07-40:30]: Ja, also, das hängt aber auch damit zusammen, dass viel Nutzer sich auch nicht daran stören. Das muss man auch sagen, es gibt so eine... Das ist so ein bisschen, hat zwei Seiten. Ich bin vielleicht zu alt, ich störe mich noch an vielem, ich möchte das nicht, vielleicht bin ich da zu tradiert in dem wie ich denke. So eine ganz junge Generation, also die meiner Kinder oder ein bisschen älter, die sagen ja und das ist doch alles ganz normal, ich kenn' das gar nicht anders! Das ist ja genau der Punkt: Wenn ich mich ja nicht daran gestört fühle, sage das ist ja schon OK wenn der das sammelt, dann ist ja meine Privatsphäre eigentlich nicht angegriffen. Ja, wo zieht man da die Grenzen. Da kann man lange glaube ich drüber nachdenken, und, das hat eher einen philosophischen Aspekt.

Also wenn wir jetzt nochmal hundert Jahre zurück denken, was wusste ein Unternehmen da über einen Menschen: Nichts – oder Alles! Also, im Dorf im Tante Emma Laden wusste sicherlich die Verkäuferin alles über ihre Kunden, weil es entsprechende Informationskanäle gab. Aber ein Dorf weiter wusste man schon wieder nichts davon. Heute ist gerade der Tante Emma Laden überall.

Ja, und da ist sicherlich der Gesetzgeber gefragt. Das Problem ist halt einfach an der Stelle ein Stück weit, die Unternehmen agieren ja weltweit, unsere Gesetzgebung schafft es maximal Europa-weit. Also, und da entstehen natürlich große Lücken, oder Ungleichgewicht. Da ist natürlich die Angst, durch die Einschränkung verlieren wir Wettbewerbsfähigkeit, uns geht es schlechter dadurch. Das möchten wir aber nicht. Wir möchten informativ selbstbestimmt sein, wir möchten aber auch ein gutes Leben führen. Und das ist ein Konflikt, und das sehe ich in den nächsten zehn Jahren nicht, dass die Welt sich mal soweit einigt, dass man sagt: Dann machen wir mal gemeinsame Sache. Da haben sicherlich die Unternehmen einen Vorteil.

R [40:30-41:02]: Und halten Sie es für wahrscheinlich, dass es aber genug, dass praktisch dieser Paradigmenwechsel in den Gedanken der Menschen, dass der in den nächsten zehn Jahren schon sehr in die Richtung geht, dass man das nicht mehr als Verletzung der Privatsphäre betrachtet? Also, dass wir dann auch sagen: Naja, ich hab ja was davon. Das gehört so.

E4 [41:02-41:30]: Das kann ich mir wohl vorstellen. Also, es gibt ja den Begriff der Digital Natives, und das ist so. Ich kann das grad beobachten, also live in einem Mikrokosmos. Und das führt zu einem ganz anderen Umgang mit den Dingen. Ob der nachher gut ist, ist eine andere Frage. Das kann ich nicht werten heute. Nicht jede Entwicklung in der Menschheit war immer gut.

R [41:31-42:23]: Dann haben wir da schon mal ganz schön viel abgedeckt. Jetzt kommen wir nochmal zurück zu den Datenmengen. Dieses Mal in Bezug auf das Netzwerk. Es gibt ja verschiedene Technologien, wie diese Daten dann übertragen werden können – Sie meinten ja auch schon im Auto ist es heutzutage gar nicht der Fall, die werden dann nur mal in der Werkstatt manuell Stück für Stück oder in Teilen ausgelesen. Können Sie sich da irgendwelche Szenarien vorstellen, dass zum Beispiel das Netzwerk gar nicht mit kommt, oder dass es Insellösungen für Einzeltechnologien gibt? Wenn man jetzt mal dieses IoT-Paradigma sieht: Überall Sensoren. Kommen dann die Netzwerke mit?

E4 [42:23-42:50]: Das ist eine gute Frage. Ich denke das Netzwerk wird sicherlich so ein bisschen eine Achilles-Ferse zunächst mal sein. Weil das Netzwerk braucht, oder ist eine große Infrastrukturmaßnahme. Wenn man jetzt mal sieht, ich weiß gar nicht wie lange ist die Diskussion hier in Berlin, WiFi hier im Stadtgebiet zu bieten, auf jeden Fall ein paar Jahre schon, oder?

R [42:50-42:51]: Auf jeden Fall schon länger!

E4 [42:51-44:22]: Auf jeden Fall schon länger, ja. Umgesetzt ist es immer noch nicht – also jetzt soll es ja langsam werden. Also das braucht relativ lange. Weil natürlich solch große Infrastrukturmaßnahmen wieder auch den Staat brauchen – das kann eigentlich kein Unternehmen leisten. Also man kann sich ja vieles denken, ein Auto, gehe ich vorwärts, wenn jede Straße so viel WiFi ein, dass sozusagen die Straße ein großer Router wird, und das Auto gar nicht mehr über drei, vier, fünf G kommuniziert, sondern direkt mit der Straße. So, das wäre jetzt ein Infrastrukturprojekt, das 30 bis 40 Jahre dauert. Und jetzt mal 30 bis 40 Jahre als Schritt, was haben wir denn schon alles, dann wird sicherlich man immer ein Stück hintendran sein und sagen, OK, da hätten wir gern [inaudible]. Also schnell voran schreiten. Also ich hätte mir vor ein paar Jahren nicht vorstellen können, dass ich jetzt zumindest in größeren, eigentlich überall wo Bebauung ist, habe ich LTE, kann eigentlich arbeiten, mehr oder weniger so, fast wie im Netz zuhause. Also wir haben sozusagen schon da auch im Netzwerk relativ starke Fortschritte gemacht, das wird es auch noch weiter geben.

R [44:23-44:55]: Und stellen Sie sich das wirklich so vor, also jetzt mit dem Zeitrahmen zehn Jahre, dass dann diese völlige Vernetzung eintritt, oder eher, dass es noch so ein bisschen langsamer geht, dass die Autos zwar erstmal alle ihre Daten sammeln, und auch Devices, diese Wearables, was auch immer es alles geben wird, und dann mehr in batches transferieren, oder wirklich diese totale Vernetzung?

E4 [44:55-46:17]: Also es wird wahrscheinlich beides passieren, auf der einen Seite glaube ich nicht, dass alle Daten, die in den Devices erzeugt werden, sozusagen eins zu eins übertragen werden. Sondern es wird so eine Vor-Aggregation, also die Devices werden intelligent sein, und ihre Daten schon ein Stück weit auswerten, und diese Auswertung übermitteln. Sodass die Datenmenge geringer wird, sodass die Netzwerklast geringer wird. Und dann an hotspots oder wie auch immer dann alles übertragen. Das kann ich mir durchaus vorstellen. Also die vollständige Vernetzung – ja – aber nicht auf den feinst-granularen Daten die da sind, sondern schon auf vor-aggregierten Daten, [inaudible], die Devices so intelligent sind das schon zu tun. Dann das Auto ist vernetzt, aber übermittelt nicht ständig seine Sensordaten, sondern wird vielleicht, wenn irgendwas außergewöhnliches passiert das direkt an die Werkstatt schicken, und dann alle Stunde mal einen Zwischenbericht von den wichtigsten Sachen. Und dann wenn es entweder in der Werkstatt, oder es ist daheim am WiFi hotspot angelangt, dann wird mal so ein batch alles übertragen. Das glaube ich, das kann ich mir vorstellen. Ja.

R [46:18-46:36]: Und auch damit gleich in Zusammenhang überhaupt das Thema der Sensoren: Halten Sie diese Prognosen für realistisch, oder eher einen Hype, dass es wirklich so schnell explodieren wird, dass wir in zehn Jahren überall Sensoren haben?

E4 [46:36-46:43]: Ich glaube da kommt ganz. Also eins ist die Energiewende. Ich glaube das... ja, zehn Jahre.

R [46:43-46:46]: Das ist natürlich weit, aber...

E4 [46:46-47:23]: Ne, es ist aber langsam auch gar nicht so viel. Ich glaube ja, dass irgendwann jeder Haushalt wird seine Stromverbräuche an eine wie auch immer geartete Stelle schicken. Das sind schon relativ viele Haushalte Deutschlandweit. Die modernen Autos werden's übermitteln, Handies tun's sowieso schon. Wearables Devices werden es auch tun. Also von daher... Ganz abgesehen von dem was jetzt in der Industrie passiert, das ist ja auch noch eine eigene Welt dann.

R [47:23-47:32]: Und so allgemein sehen Sie dann die meisten Potenziale eher im, sagen wir jetzt mal, was die Endkonsumenten betrifft, oder eher in der Industrie?

E4 [47:32-47:59]: Beides. Das wird sich die Waage halten. Also im Consumer-Bereich, das sehen wir ja, da wird's viel geben weil es eben auch viele Anwendungen gibt, das ist ein großer Markt. Da wird viel voran gehen. Und Industrie-Bereich wird eben in die Richtung der Automatisierung weiter gehen. Und Effizienzsteigerung.

R [48:00-48:43]: Und mal global gesehen, könnten Sie sich vorstellen, dass die Bereiche, die jetzt noch sehr weit hinterher sind, also vor allem Entwicklungs- und Schwellenländer, dass die dann direkt aufholen, oder, dass die Schere noch größer wird? Also es gibt ja schon Projekte, zum Beispiel für die Vernetzung dieses Google Loon Projekt, wo die Ballons für WiFi aufstellen. Andererseits habe ich neulich auch gelesen, zum Beispiel in Afrika, das Netzwerk ist so nicht vorhanden, dass Google einen SMS-Service eingerichtet hat.

E4 [48:43-49:44]: Also ich bin grad am überlegen, weil da gibt es ja verschiedene Dinge. Also jedes Land was heute hinterher hinkt, da gibt es eine Grundvoraussetzung: Das ist eine Befriedung. Wenn wir dann schauen welche Regionen der Welt hinken da hinterher, das sind ausgerechnet die, wo relativ viel Krieg und Unfrieden herrscht. Und solange das herrscht werden die Länder keine Chance haben, das ist einfach eine Grundvoraussetzung. Wenn das mal eintrifft, dann werden die natürlich nicht die Evolution durchmachen, die wir durchgemacht haben. Sondern die werden dann schon mit den neuen Technologien starten. Und werden dann hoffentlich ein Stück weit die Lücke schließen, ob sie die dann komplett schließen ist immer eine gute Frage. Das hängt natürlich auch von den anderen Ländern ab, in wie weit die das zulassen.

R [49:45-49:49]: Können Sie sich irgendwelche großen potenziellen Einflussfaktoren die diese Entwicklung aufhalten, also vor allem in Bezug auf Verbreitung von Sensoren und auch Vernetzung?

E4 [50:00-50:41]: Ein Meteoriten-Einschlag. Nein, ich glaube wirklich, es kann... Also, globale Katastrophen, ein Weltkrieg, ein Meteoriten-Einschlag. Das ist etwas, was uns so zurück werfen könnte, dass wir das nicht... Dass es nicht weiter geht. Aber in einer normalen Entwicklung: wir können das ja alles schon. Und dann wird die anhalten. Also wenn wir keine globale Katastrophe, dann wird es so weiter gehen.

R [50:41-50:52]: Ich würde sagen, das ist ein ganz guter Abschluss. Dann sind wir jetzt eigentlich durch alle Themen durch.

[50:53-53:31]: *[Closing conversational remarks.]*

Transcript E5

Date: 29.01.2016

Duration: 31:53 [mm:ss]

Modus: face-to-face interview

Interviewee:

Expert 5 [E5]: CEO and Co-Founder of a German start-up, providing an enterprise application based on a relational in-memory database; previously obtained a PhD in the field of in-memory database systems

Interviewers:

Julia Reichelt, Researcher [R]

[00:00-00:52]: [introductory remarks]

R [00:53-01:23]: Genau, also das erste ist die Echtzeit, überhaupt die Nachfrage nach Echtzeit-Analyse, und Echtzeit... Beschleunigung von Services und so weiter. Und da habe ich mir zunächst überlegt: Es ist interessant, wer braucht sowas überhaupt? Braucht das jedes Unternehmen irgendwann? Oder ist das doch eher so etwas wo man sagt: OK, nicht jeder wird das...

E5 [01:23-02:36]: Also vielleicht... als erste Anmerkung: Der Begriff Echtzeit ist immer von der Verwendung ist immer ein bisschen mehrdeutig, zweideutig. Echtzeit heißt ja eigentlich, dass man Daten, so wie sie gerade vorliegen, also die letzten Änderungen, die gerade im Geschäft stattgefunden haben, dass man die nutzt zur Analyse. Das heißt aber weniger, und das ist glaub ich hier der Kernpunkt, dass man die Ergebnisse so schnell wie möglich bekommt. Also den Echtzeit-Begriff da vielleicht trennen.

Zu den Fragen: Also ich bin nach wie vor der Meinung, dass es eine hohe Relevanz hat für die Unternehmen, einfach weil die Unternehmenskomplexität auf der einen Seite gestiegen ist, und zum anderen auch die Komplexität, die Volatilität in der Umgebung. Und so müssen wir... viel eher, viel stärker in der Lage sein, da schnell agil zu reagieren auf Veränderungen im Umfeld des Unternehmens.

R [02:36-02:40]: Wird sich das auch weiter steigern?

E5 [02:38-03:55]: Ja, ja, auf jede Fall. Das spannende bei der ganzen Sache ist ja die, dass zum einen die Datenvolumina immer stärker anwachsen, und zum anderen natürlich der Bedarf auch aus diesen Daten, ja zu schöpfen, also da Informationen heraus zu ziehen und... Das wird weiter steigen. Also man sagt ja nicht ohne Grund, dass data is the new gold, dass da unheimlich viel Wert drin steckt. Und es gibt ja unzählige von Unternehmen, die sich darauf spezialisieren diese Schätze zu heben. Und das natürlich so attraktiv wie möglich zu machen ist der Punkt: Echtzeit, oder schnelle Verarbeitung von Daten ist da von zentralem Interesse. Und das wird sich auch in den nächsten zehn Jahren weiter... Also, da kannst du auch Gartner Studien mal zitieren dieses Datenwachstum, Verdopplung alle zwölf Monate, alle achtzehn Monate...

R [03:56-04:31]: Der Interviewpartner von heute früh meinte, dass er das in manchen Bereichen noch nicht so viel sieht, und zwar – jetzt muss ich eben gucken – genau, dass diese predictive analytics, hat er noch nicht so viel beobachtet im Markt. Wie beurteilst du das? Dass das also auch steigt? Oder ist das gerade ein ganz neues Feld, das gerade erst im Kommen ist?

E5 [04:31-05:13]: Naja, klar, also das ist ein... Also predictive heißt ja in die Zukunft schauen, intelligente Vorhersagen zu machen. Und dieses ganze Feld wird halt auch dahingehend interessanter, weil die Technologien auch ausgereifter sind, und immer mehr Datenquellen da auch, und eine Vielzahl von Daten unterschiedlichster Art heran, herbeigezogen werden. Analysen so genau wie möglich zu machen. Finde ich ist vielleicht ein noch viel größerer Trend als das Thema Analytics, also Analyse der Vergangenheit.

R [05:14-05:19]: Kennst du irgendwelche Bereiche, wo das besonders relevant ist?

E5 [05:20-05:54]: Ein Begriff: Predictive Maintenance zum Beispiel, wo man halt besser vorhersagen will wann Maschinen oder Bauteile gewartet werden müssten. Also bei Flugzeugen, oder unterschiedlichste Maschinen: Es da so spät wie möglich zu machen, aber so früh wie nötig, dass man keine Ausfälle hat. Das ist dann halt auch die direkte Kostenersparnis, die da auch her kommt.

R [05:55-06:05]: Also eher auch im industriellen Bereich, als jetzt zum Beispiel Consumer-Applications?

E5 [06:06-06:50]: Mhm. Also da ist vielleicht der direkte monetäre Wert dahinter einfach viel größer. Im Consumer-Bereich ist es ja eher so ein Usability-Aspekt, dass wenn man halt nicht Minuten oder Stunden warten will, bis da ein Ergebnis kommt. Also stell dir vor du willst bei Amazon shoppen, und dieses Kunden-Kauf-Auch kommt erst nach einer Minute, dann nützt das ja sicher gar nix. Oder Google, Google ist das beste Beispiel für dieses ganze Realtime und Echtzeit. Ist ja einer der Gründe des Erfolgs von denen.

Sollen wir die Kosten auch gleich noch dazu besprechen?

R [06:48-06:54]: Gut. Genau, das ist der nächste Abschnitt.

E5 [06:54-07:01]: Vielleicht eine Frage vorab: Kennst du dieses Buch In-Memory Revolution von Bernd Leukert und Hasso Plattner?

R [07:01-07:10]: Ich habe eins von denen... Das ist relativ neu von 2015.

E5 [07:10-07:53]: Genau. Lies dir das mal durch, weil da sind auch viele von diesen Fragen, also jetzt zum Beispiel der Kosten-Aspekt... Also klar sind die Systeme teuer für den Enterprise-Bereich. Aber: der TCO, also die Gesamtkosten des Systems sind häufig viel geringer als herkömmliche Datenbanksysteme. Das zum einen. Und zum anderen hast du natürlich, wenn du so ein zentralen Business Value wie jetzt predictive maintenance brauchst, kannst deine Wartungskosten um 300% senken, dann hast du die Investition auch ganz schnell raus.

R [07:54-08:16]: Und zum Beispiel sowohl SAP HANA als auch andere relationale in-memory Datenbanken [RIMDB] sind ja meistens auf Lizenzsystemen basiert. Könntest du dir vorstellen, dass es auch was anderes gibt? Zum Beispiel, das ist auch in enger Verbindung mit dem nächsten Block, dass sich dann open-source Lösungen entwickeln?

E5 [08:17-08:19]: Ja gibt's ja. Es gibt open-source Lösungen.

R [08:19-08:21]: Auch für RIMDB?

E5 [08:21-08:22]: Ja.

R [08:22-08:25]: Weißt du da welche?

E5 [08:26-09:07]: Wie heißen die... MonetDB zum Beispiel... Das ist... Da gibt's eine ganze Reihe von, fällt mir jetzt nicht ein. Ist halt immer die Frage: Produktiver Einsatz, und dann auch Kosten, Anpassung, etc. Ich glaub der Kostenfaktor, den kann man hier echt aus meiner Sicht vernachlässigen, und da kannst du auch gern mal in das Buch schauen. Da gibt's glaube ich ein, zwei Kapitel zu dem Thema.

R [09:07-09:53]: In vorherigen Interviews habe ich schon gehört, dass es sich für viele Kunden noch nicht so richtig lohnt. Habt ihr diese Erfahrung dann nicht gemacht? Im Vergleich zu anderen Datenbanken, die mehr disk-basiert sind, und praktisch auch dadurch immer schneller... Oder zum Beispiel durch Kopplung mit Hadoop-Systemen immer schneller werden? Ich hatte da zum Beispiel auch jemanden, der Beratung und Implementierung von solchen Systemen macht, interviewt, und der meinte für seine Kunden ist das einfach oft zu teuer. Und zwar nicht nur Mittelständler, sondern auch wirklich größere Firmen.

E5 [09:54-10:23]: Ja klar. Es gibt immer die... Es ist immer eine Frage des Einsatzes und des Business Cases, den man damit lösen kann. Kann man so pauschal nicht sagen, das ist günstig, oder das ist teuer. Also SAP HANA hat natürlich einen sehr hohen Price-Point, in diesem Punkt gebe ich dir Recht.

R [10:24-10:55]: Und in Bezug auf die wachsende Datenmenge – das ist ja immer... Also die ganzen Prognosen, auch zum Beispiel von Gartner, sagen es wird unfassbar explodieren wie viele Daten da ausgewertet werden müssen. Meinst du, dass sich das in irgendeiner Form auf den Kostenpunkt einwirkt? Oder wird... Es ist ja immer die Frage: Die hardware wird zwar günstiger, aber die Datenmenge steigt.

E5 [10:56-12:21]: Ich... Ja... Im Vergleich zueinander... Schwer zu sagen. Also die hardware-Kosten kann man glaube ich auch vernachlässigen... Es ist immer eine Frage des Use Cases, und auch wo du jetzt Hadoop, und Spark, und die ganzen anderen Technologien in den Raum geworfen hast. Das sind ja leicht andere Use Cases,

die damit bearbeitet werden. Weniger dieses Echtzeit-Echtzeit, also ich klick in meinem BI-Dashboard rum und will sofort sehen was steckt dahinter, sondern das sind eher so Jobs, die ein paar Sekunden länger dauern können. Vielleicht doch mehr unstrukturierte Daten bearbeiten als jetzt hier eine RIMDB, das ist ja eher für ein ERP-System. Und da sind ja auch die... Die Datenmengen und das Datenwachstum nicht ganz so stark wie im unstrukturierten Bereich, also doch etwas begrenzter.

R [12:22-12:44]: Wobei sich jetzt auch... Also wenn man jetzt mal Produktionsprozesse betrachtet, sowas kann ja auch mit relationalen Datenbanken verwendet werden, und auch da sind Prognosen, dass sich jetzt – mit dem Schlagwort Industrie 4.0 – die Sensorenmenge wachsen wird.

E5 [12:45-12:58]: Ja, das stimmt. Das ist richtig, ja, und das wird wahrscheinlich dann der Punkt.

R [12:59-13:31]: Was ich versuche hier heraus zu finden ist, dass es nicht nur eine lineare Entwicklung darstellen kann, sondern zum Beispiel Optionen wie das für mittelständische, kleinere Unternehmen immer attraktiver wird, weil die sich zum Teil auch keine... Die haben zum Teil einfach nicht die Kapazitäten sich so eine in-memory Datenbank zu leisten. Oder würdest du dann sagen, OK, für die ist das einfach nichts? Die Kosten...

E5 [13:31-14:17]: Ja da vielleicht noch einmal zum Thema Lizenz, ich weiß nicht ob du das... Lizenzmodelle: Es gibt ja auch Cloud-Modelle, wo man ohne großes Investment gegen eine monatliche Gebühr sich so ein System mieten kann. Ich glaube das hängt auch weniger von der Größe des Unternehmens ab, als von der Datenzentriertheit, oder Datenintensivität, die dort im Unternehmen anfällt. Wie viel Wert in den Daten steckt, und wie schnell, oder was der Mehrwert ist, diese Daten auch schnell zu analysieren, auszuwerten.

R [14:18-14:32]: Glaubst du, dass sich across all industries der Mehrwert von Daten steigern wird in den Unternehmen mit Zeithorizont die nächsten zehn Jahre, oder ist das doch eher schleppend?

E5 [14:33-14:49]: Schwierig zu beantworten, aber ich denk mal grundsätzlich schon, ja. [inaudible]

R [14:50-15:18]: Okay, dann können wir gleich zum dritten Punkt gehen, und zwar da geht's noch einmal konkret um diese open-source Modelle. Du sagtest ja schon, es gibt schon welche im Markt. Könntest du dir auch vorstellen, dass es passieren kann, dass zum Beispiel, dass SAP, Oracle, IBM gedrungen werden in diese Richtung zu gehen?

E5 [15:18-15:19]: Warum, warum? Warum sollten die das machen?

R [15:19-15:38]: Das ist ja anscheinend auch in anderen Bereichen von Software passiert, dass es dann eher in open-source ging im Nachhinein, und die Business Models eher um die zusätzlichen Services als Add-On...

E5 [15:38-16:19]: Also das glaube ich jetzt nicht bei den großen Playern. Würde keinen Sinn machen, also... SAP hat ja jetzt auch nicht das ERP-System, dass sie vor, oder seit 40 Jahren entwickeln, als open-source Lösung bereitgestellt. Es gibt eher dann andere Anbieter, die anfangen etwas ähnliches nachzubauen, um das dann in open-source Manier voran zu treiben. Aber dass da SAP, Oracle, oder EXASOL, oder wie sie alle heißen das machen, das macht aus kommerzieller Sicht keinen Sinn. Also aus meiner Perspektive jetzt.

R [16:20-16:36]: Würdest du dann zum Beispiel... OK, würdest du SAP HANA eher in Verbindung mit ERP-Systemen sehen, also dass der Mehrwert durch die Verbindung mit dem ERP-System entsteht?

E5 [16:37-17:37]: Vielleicht da noch ein Punkt: Es bringt halt oft wenig, wenn man sagt, ich tausche jetzt die Datenbank aus, und packe da jetzt die SAP HANA drunter, und es wird alles schneller und viel besser, und... dem Unternehmen geht's gut, sondern es ist in vielen Fällen auch die Anwendung, die angepasst werden muss. Und deshalb wird ja auch bei der SAP ja gerade dieses SAP... ERP-S4-HANA gebaut, was ja genau das macht: Die Anwendung komplett umstrukturieren. Und dann können die Vorteile der schnellen Datenbank genutzt werden. Neben dem gibt es halt viele, viele, neue Anwendungen, die jetzt erst überhaupt möglich werden mit den schnellen Datenbanken.

R [17:37-17:39]: Zum Beispiel?

E5 [17:40-18:56]: Ein paar Beispiele, also predictive maintenance, ist ja jetzt so das, was schon [inaudible], und hängt natürlich auch mit dem Thema Industrie 4.0 zusammen, soweit die Daten verfügbar sind, und es gibt ein paar andere Beispiele. Also wir haben am Lehrstuhl vorher so Sachen gemacht wie cross-selling, dass bei

Amazon halt zum Beispiel funktioniert, solche Geschichten, oder interaktive Dashboards, die... Also da gibt's... Da würde ich dir auch nochmal dieses Buch empfehlen, da sind auch nochmal viele von diesen Anwendungen drin. Also Thema Simulation, Naturkatastrophen, Einfluss auf Versicherungsschäden, die dann bei einer Flut zum Beispiel bezahlt werden müssten. Also die Bandbreite ist da doch schon sehr, sehr groß.

R [18:57-19:53]: [inaudible] den Punkt. Dann können wir auch weiter gehen zu Block vier. Da geht es um die gesetzlichen Regulierungen. Und vor allem wichtig wäre für mich zu wissen wie... was kann passieren, dass die Datensammlung stark eingeschränkt wird, oder was auch die Regierungen auswirken können, um die Potenziale optimal zu nutzen. Man kann natürlich auch super viel... Klar, man muss Daten irgendwie schützen, aber andererseits kann man auch total viel machen, extreme Effizienzsteigerung durch Verwendung von Daten.

E5 [19:54-19:55]: Von gesetzlicher Seite?

R [19:55-20:23]: Genau, also wie sieht es aus von gesetzlicher Seite. Es gibt ja zum Beispiel jetzt diese EU-Richtlinie für Privatsphärenschutz. Was könnten da für Entwicklungen passieren, dass zum Beispiel die Entwicklung des Internet of Things [IoT], also der vielen Sensoren, und des starken Datenvolumens eingeschränkt wird, oder um die Potenziale noch besser zu nutzen, als man es heutzutage macht.

E5 [20:24-20:39]: Ich wüsste jetzt nicht was dafür sprechen sollte das einzuschränken. Weil so ein Sensor an so einer Bohrmaschine hat jetzt wenig Privatsphäre.

R [20:40-21:12]: Zum Beispiel, wenn man jetzt an self-driving cars denkt, sind da natürlich ein Haufen Sensoren drin. Und dann ist der in irgendeiner Form ans Netz angeschlossen, und theoretisch weiß dann.... Kann dann zu jeder Zeit bestimmt werden, wo die Person, die in diesem Auto gerade herum fährt ist, und mit was für einem Auto sie unterwegs ist. Beispielsweise, dass man personenbezogene Daten, die...

E5 [21:11-21:42]: Glaube ich nicht, dass das jetzt ein Thema ist in dem Bereich. Also es hat wenig mit in-memory Datenbanken zu tun. Es ist eher ein generelles Problem, und ich meine wir haben jetzt irgendwie alle ein iPhone, und da weiß Google auch wo wir sind, wo wir sitzen, was wir gerade machen, und hat auch keiner ein Problem dabei.

R [21:43-22:05]: Also du glaubst nicht, dass da... Es geht nicht nur darum, wie es in in-memory Datenbanken ausgewertet wird. Sondern die Grundlage ist ja praktisch die Datensammlung. Also dass ich ja selber... Ich muss ja erstmal die Daten sammeln...

E5 [22:05-22:07]: Ist die jetzt nur für Deutschland, die Analyse, oder ist die weltweit?

R [22:07-22:09]: Ne, weltweit. Wobei...

E5 [22:08-22:48]: Dann würde ich dem Thema keine größere Beachtung schenken. Den Amis ist es völlig egal, dem Rest der Welt glaube ich auch. Solange der Mehrwert gesehen wird, wird's immer Leute geben, die bereit sind denn auch die andere Seite der Medaille zu ertragen. Ist ja bei Kreditkarten genauso. Weiß ja auch jeder wo man wann bezahlt hat. Und wir machen's trotzdem, warum, weil es einfach convenient ist.

R [22:49-23:59]: Genau, [inaudible]. Und dann auch praktisch als Grundlage, wieder bezüglich des Big Data Themas, und IoT. Das Netzwerk muss natürlich erstmal da sein, um überhaupt diese ganzen Daten zu sammeln. Wenn man da jetzt eine eigene Fabrikhalle anschaut, das ist das klar, dann kann das jede Firma selber machen. Aber jetzt allgemein diese prognostizierte totale Vernetzung: Da muss jetzt erstmal noch einiges geschehen, damit das Netzwerk da mit kommt. Wie stellst du dir das vor? Glaubst du es wird... Zum Beispiel wenn du dir jetzt Berlin anguckst: OpenWiFi ist schon ewig lang in Diskussion, und es passiert nur sehr langsam, dass es wirklich auch umgesetzt wird. Und auch in ganz Deutschland ist das Handynet – in Großstädten natürlich schon – aber in ländlichen Regionen nicht so gut. Und wenn die Datenmenge so explosionsartig steigen würde, wie vorher gesagt, hätte das Netzwerk natürlich ein Problem.

E5 [24:00-24:37]: Also verstehe ich nicht ganz die Problematik in dem Punkt. Das was du jetzt gerade angesprochen hast, WiFi, Handynetze, das ist ja dann nur die Schnittstelle zum Endverbraucher. Viel entscheidender sind ja die Backbones die dann quasi das Internet zusammen halten. Die Analysen werden dann ja eh zentralisiert, wobei sie auch zum Teil dezentralisiert durchgeführt werden. Da ist glaube ich das Netzwerk das geringste Problem.

R [24:38-24:50]: Du gehst also davon aus, dass das Netzwerk mit dieser extremen ansteigenden Sensoren-Menge auch dann zurecht kommt?

E5 [24:50-15:30]: Grundsätzlich denk ich mal... Also ich bin da kein Experte in dem Bereich. Keine Ahnung. Also da gab's ja auch so ein paar grundlegende Schritte, also IPV6, die dann in den letzten Jahren gelegt wurden. Das ist alles... Also musst du dir mal ein paar Studien anschauen, da gibt's bestimmt Expertenmeinungen, die da mehr relevant sind als meine.

R [25:30-26:00]: Die Grundidee ist halt: Klar, real-time, schön und gut, aber wenn man jetzt wirklich real-time Analysen aus den ganzen vielen Sensoren macht, die irgendwo dispers verteilt sind, dann muss natürlich auch das Netzwerk diese ganzen Informationen sammeln können, und könnte natürlich ein bottleneck entstehen, dadurch, dass das Netzwerk das eben nicht packt.

E5 [26:01-26:11]: Möglich. Kann ich schlecht drauf antworten. Ich denke es ist kein Problem, aber...

R [26:12-26:14]: Dann sind wir schon beim letzten Punkt.

E5 [26:14-26:46]: Also, vielleicht doch. Eine Antwort. Es gibt ja Moore's Law und irgendwelche anderen Gesetzmäßigkeiten in der Entwicklung von Technologien, und da ist zum Beispiel das Thema, dass die Netzwerkgeschwindigkeit, wenn ich das Recht wiedergebe, sich stärker entwickelt hat als zum Beispiel große Prozessoren-Geschwindigkeiten, Rechenkapazitäten.

R [26:47-26:52]: OK, dementsprechend der Trend geht wahrscheinlich so weiter.

E5 [26:52-26:53]: Ja.

R [26:54-27:18]: Dann zum letzten Punkt, da geht es konkret um die Sensoren. Hältst du diese Prognosen eher für einen Hype? Oder denkst du, das wird schon auch so zutreffen? Wie beispielsweise Gartner das vorhersagen mit was weiß ich wie vielen Trillionen an devices by 2030.

E5 [27:29-28:01]: Also ich kann mir das schon vorstellen, dass das weiter wächst. Einfach weil ich den Wert sehe: Über das Messen von Zuständen in der Industrie, wo auch immer, smart home... Und wenn ich überlege es gibt ja auch die wristbands, die ganzen fitness-tracker, und das sind ja alles Datenpunkte, die gesammelt werden. Und warum macht man das? Einfach weil man sich über die Analyse Optimierungspotenzial verschafft – oder verspricht. Das wird weiter wachsen, denk ich mal.

R [28:02]: Und kannst du dir ein Negativ-Szenario vorstellen, warum das nicht passieren könnte? Oder zumindest, dass es stagniert? [inaudible]

E5 [28:17-28:52]: Ne, also ich bin Optimist, ich denke das passt. Es sei denn, dass es tatsächlich so ein Privacy-Thema wird, dass man... Also gerade jetzt wenn man über diese Fitness-Tracking, oder generell... Steckt in der eigenen Person... Wenn's darum geht, dass man dann einfach die Nachteile überwiegen werden.

R [28:52-28:59]: Genau, das ist ja in Deutschland immer dieses Stichwort Der Gläserne Bürger....

E5 [28:59-28:12]: Also ich glaube da sind wir schon sehr gläsern. Und schadet jetzt nicht wirklich. Also aus meiner Sicht.

R [28:13-29:26]: OK. Das nächste haben wir eigentlich schon besprochen... Die anderen Bereiche... Dann war's das auch schon.

E5 [29:26-29:28]: Dann war's das.

[29:28-31:53] [thanking, closing remarks, goodbye]

Transcript E6, E7

Date: 04.02.2016

Duration: 52:35 [mm:ss]

Modus: face-to-face interview

Interviewee:

Expert 6 [E6]: Research Assistant and PhD Student in the field of in-memory data databases and related topics

Expert 7 [E7]: Research Assistant and PhD Candidate in the field of in-memory database technology and data streaming

Interviewer:

Julia Reichelt, Researcher [R]

[00:00-01:12] [Introductory remarks]

R [01:13-01:21] Vielleicht könntet ihr nochmal jeder nochmal kurz erzählen, inwiefern ihr euch mit relationalen in-memory Datenbanken beschäftigt [RIMDB].

E6 [01:22-02:17]: Ich beschäftige mit RIMDB im Moment im Rahmen meiner Forschung eher weniger, ich bin eher im Bereich e-learning unterwegs, habe allerdings natürlich hier die Vorlesungen alle als Student damals gehört und seit dem ich eben wissenschaftlicher Mitarbeiter bin die Vorlesungen mit vorbereitet, konzipiert, also die Folien mit aufbereitet, die Online-Vorlesungen mit geleitet und mit gestaltet und bin da sozusagen in der Lehre mit vorne dabei, und beschäftige mich aus diesem Grund quasi mit dem Lehrmaterial, mit den Technologien dahinter. Ein paar Anwendungsfälle hatte ich natürlich auch mal wenn ich etwas implementiert hab auf der HANA. Meine Forschung ist es aber akut nicht.

E7 [02:19-02:40]: Ja ich... Bei mir ist es im Prinzip ähnlich, also ich forsche nicht an RIMDB, aber ich war auch Nutzer ein paar Jahre lang, und jetzt immer noch so ein bisschen. Ich habe die benutzt um da Dinge zu modellieren, zu entwickeln. Aber selbst an der Datenbank [DB] habe ich nicht gearbeitet.

E6 [02:39-02:42]: Aber du benutzt sie weiterhin, oder? Für die IoT-Daten?

E7 [02:42-02:48]: Ja schon, aber ist nicht Fokus oder zentrales Element.

R [02:48-02:53]: In was für einem Rahmen von IoT-Daten ganz grob?

E7 [02:53-03:42]: Also ich beschäftige mich mit der... mit Data-Stream-Processing Systemen, und da in dem Rahmen ist es ja manchmal auch der Fall, dass man Data-Streams anreichert mit historischen Daten beispielsweise, und da könnte man dann beispielsweise auch auf RIMDB zugreifen um da entsprechend auch Daten zu lesen. Und ich hab mir auch die Data-Stream-Verarbeitungs-komponente von einer RIMDB bisschen näher angeguckt. Und damit etwas gemacht, also spezielle von HANA, da gibt's auch eine Streaming-Komponente.

R [03:43-04:15] [*brief pause*] Gut, gleich zum ersten Themenblock: Da geht es um die Nachfrage von Echtzeit-Reaktionen beziehungsweise Echtzeit-Analyse und insgesamt die Beschleunigung von Services. Und da war so der erste Gedanke: Wer braucht sowas überhaupt? Also für wen ist das relevant heute, und wird in den nächsten Jahren immer relevanter?

E6 [04:16-06:43]: Ich kann wieder anfangen. Das was ich jetzt nenne ist das [inaudible] wahrscheinlich jegliches Marketing-Material auch hergibt, das ist ja eine relativ allgemeine Fragestellung. Also es wird halt bei allen unternehmenskritischen Entscheidungsprozessen nach dem Sinne gebraucht werden, also wenn ich jetzt das ganz klassische Beispiel habe: ein Board-Meeting, und die wollen eben schauen, OK, wie läuft denn unser Geschäft in Asien? Dann möchte man da vielleicht eben ein bisschen tiefer rein gucken, und auch eine bestimmte Sparte vielleicht selektieren, also wenn ich jetzt einen Hersteller für, weiß ich nicht, Werkzeuge bin, dann möchte ich mir vielleicht gerade mal nur Gartenwerkzeuge anschauen, und dann brauche ich ja eben diese Echtzeit-Informationen, und will vielleicht auch schauen, OK, hängen da Trends vom Wetter ab, dass ich gerade bei besonders viel Sonnenschein mehr Verkauf von Rasenmähern habe, oder irgendwie sowas.

Also das sind ja eben die klassischen Dinge, die ich jetzt sozusagen kenne, weil wir sie eben auch in der Lehre verwenden, beziehungsweise in den Online-Kursen als Motivatoren quasi: Warum braucht man sowas, wenn eben oft auch die Antwort ist: wenn ich komplexe Analysen möchte, dann kann ich da auch einen Tag drauf warten. Das ist ja eben das, was man sozusagen meistens hört, und dann bringen wir eben auch Beispiele, dass man vielleicht in einem Board-Meeting alle einen Tag warten lassen möchte, weil die Leute eben auch viel Geld kosten.

Das andere ist natürlich zur Prozessunterstützung, also wenn ich natürlich Dinge verkaufe, dann will ich vielleicht die Kassenkraft, oder wer auch immer den Verkauf vielleicht vornimmt, kann natürlich sagen... Im Supermarkt wird man das jetzt eher nicht haben, OK, wir verkaufen die Ware, weil ich krieg ja auch mein Geld direkt. Aber wenn ich jetzt Bestellungen aufgabe von Unternehmen zu Unternehmen, und da jetzt irgendwie ein Geschäftsvolumen von \$200,000 habe, dann möchte ich ja vielleicht vorher wissen wie ist denn der Zahlstand der bisherigen offenen Aufträge, die müssen ja auch nicht auf einen Schlag überwiesen werden, sondern die kommen ja dann vielleicht auch gestückelt, und wenn die dann im Rückstand sind will ich vielleicht keinen zusätzlichen Rabatt geben, und wenn ich weiß das ist ein sehr guter Kunde, dann kann ich da sozusagen auch vielleicht live dann irgendwie den Rabatt-Rahmen anpassen, andere Möglichkeiten geben, quasi zur Verkaufsunterstützung die Informationen nutzen. Das sind so zwei Beispiele, die ich schon mal nennen könnte, die halt aber auch schon, wie gesagt, sowohl im Marketing als auch in der Lehre benutzt werden als Motivatoren.

E7 [06:43-07:15]: Ja, also sehe ich genauso. Ich denke wenn man besonders aktuelle Daten möchte, zum Beispiel keine Stunde auf eine Abfrage warten möchte, weil da haben sich die Daten [inaudible] schon geändert, oder wenn man auf besonders schnell sich verändernde Zustände oder so reagieren möchte. So ganz allgemein sind denk ich Dinge die in vielen Bereichen eine Rolle spielen können und relevant sein können. Also sicherlich beeinflusst das jetzt Geld oder andere Sachen die wichtig sind.

E6 [07:16-07:54]: Beziehungsweise bei Stream-Daten schätze ich ja, dass es nur so bedingt... Also, dass es da nur das von der Hauptspeicher-DB rein kommt... Weil, also korrigier mich wenn ich da falsch liege, aber ich würde sagen auch für Stream-Daten reicht die Zeit dir dann sozusagen nicht, wenn du in einer Hundertstel- oder Tausendstel-Sekunde etwas analysieren möchtest, dann hast du halt dummerweise halt auch nicht eine Zehntel-Sekunde Zeit, auch wenn die jetzt vielleicht in einer Zehntel-Sekunde antworten würde, weiß ich nicht, ob sozusagen beim Stream-Processing für kurzfristige Anfragen das da mit rein... Also weiß ich nicht, wie da der Stand ist? Würde mich persönlich jetzt auch mal interessieren.

E7 [07:54-07:55]: Ich glaub ich hab dich noch nicht ganz verstanden.

E6 [07:56-08:46]: Du hast im Prinzip... Du kriegst einen Datenstrom rein, die Antwort ist, also ich weiß zum Beispiel nicht was der Strom ist, du kriegst irgendwelche Sensorwerte rein von einer Flugzeugturbine oder sonst was. Du kriegst du ja in einer massiven Geschwindigkeit und du suchst nach Anomalien. Und wenn du da jetzt sozusagen zusätzlich irgendwie Informationen aus einer Hauptspeicherdatenbank bräuchtest – ich hab grade kein Beispiel zur Hand wo das Sinn machen würde – dann brauche ich ja wahrscheinlich diese Anomalie-Detektion brauche ich ja wahrscheinlich in einer Hundertstel-Sekunde, und wenn meine in-memory DB halt nur in einer zehntel Sekunde antwortet, was ja auch immer noch schnell ist, dann reicht das ja sozusagen nicht. Die Frage ist sozusagen: welche Daten aus der schnellen DB kann ich verwenden oder bleiben so lange stabil, dass das auch für den Anwendungsfall Echtzeit-Systeme und Streaming-DB...?

E7 [08:46-08:53]: Klar, also im Endeffekt ist es nur so schnell wie die langsamste Komponente, oder man muss das cachen oder parallelisieren...

E6 [08:53-09:00] ...oder halt vorbereiten, man weiß jetzt, dieser Wert gilt fünf Minuten oder so.

E7 [09:00-09:05] Das ist halt stark abhängig vom use case den man hat.

R [09:06-09:20]: Also gerade bei so Sensoren würdet ihr dann eher sagen, die in-memory DB, wenn man jetzt wirklich in sehr kleinen Bruchteilsekunden Infos braucht, sind die doch noch ein bisschen langsam?

E6 [09:20-09:45]: Ne, die sind eigentlich nicht zu langsam, die Frage ist sozusagen, je nachdem, also auch jede andere DB, wenn ich viele viele Daten verarbeiten muss, dann habe ich da auch einfach irgendwo eine Begrenzung wie viele Daten ich angucken kann, bis ich mein Zeitlimit erreicht habe. Und die Streaming-DB

gucken sich, also wenn ich es richtig in Erinnerung habe, immer nur bestimmte Fenster an von Daten die rein kommen, und nicht den Gesamtdatenbestand den sie sozusagen durchgeschleust haben. Oder?

E7 [09:46-09:49]: Ja, also es wird...

E6 [09:49-09:56]: Und es ist halt ein anderer Anwendungsfall. Die in-memory DB kann alle Daten angucken um eine vollständige Antwort zu liefern.

E7 [09:56-10:00]: Es ist halt ein anderes System einfach.

R [10:00-10:08]: Das kommt dann einfach auf den Kontext an ob man auch wirklich alle Daten anschauen muss, oder nur bestimmte Teile ausreichen um die Aussage zu treffen.

E7 [10:08-10:30]: Das ist ja auch überhaupt die Frage ob man dann alle Daten überhaupt speichern muss oder möchte, das ist ja dann auch wieder eine Kostenfrage – je mehr Daten, desto mehr Speicherplatz, die auch hardware brauchen, und die auch nicht die günstigste ist in der Art. Viel schichtigere Fragestellung, die wahrscheinlich individuell beantwortet werden muss von Unternehmen zu Unternehmen, use case zu use case.

R [10:31-10:5]: Wie glaubt ihr, dass sich das entwickelt über die nächsten zehn Jahre? Also wird das dann... Haben wir praktisch jetzt schon einen Stand erreicht, dass die Nachfrage an Echtzeit-Reaktionen sehr hoch ist, oder wird das noch sehr stark steigen? *[brief pause]* Einfach mal eine grobe Einschätzung, das muss jetzt natürlich nicht ganz...

E6 [10:55-11:11]: Ich würde vermuten, dass wir uns weiter entwickeln auf eine Standard-Technik, und immer mehr in Echtzeit möglich, und dass es auch immer mehr State of the Art wird, und verlangt wird. Und was war die erste Frage?

R [11:11-11:15]: Wie sich das jetzt so über die nächsten zehn Jahre...

E 6[11:15-11:31]: Ich denke tendenziell werden wir nicht langsamer sondern schneller. Ich denke das haben wir in den letzten Jahren auch schon gesehen. Ich denke das wird sich nicht ändern, auch wenn Datenmengen eventuell steigen. Ich denke es wird mehr so in den Bereich cloud gehen, verteilte Systeme und so weiter.

R [11:32-11:50] Und jetzt von Unternehmensseite, glaubt ihr, dass auch wirklich, dass da die Nachfrage sehr stark steigen wird? Oder sind die noch eher am Anfang der Realisation: Wir brauchen das jetzt mal. Oder haben das schon sehr viele implementiert? Was sind da eure Erfahrungen?

E7 [11:50-12:27]: Also Veränderungen der IT-Landschaft sind natürlich immer mit Gefahren und Kosten und so weiter verbunden. Das ist natürlich immer ein schwieriger Punkt. Aber auf der anderen Seite darf man natürlich auch nicht komplett abgehängt werden auf IT-Seite, weil das natürlich auch Anwendungen auf der Business-Seite zur Folge hat, und dass sind natürlich höchstwahrscheinlich Unternehmen, auch wenn sie vielleicht nicht immer zwingend gewillt sind, teilweise gezwungen da auch zumindest stückweise sich neuen Entwicklungen anzupassen. Und neue Unternehmen, Start-Ups, und so weiter die haben es da natürlich deutlich einfacher.

E6 [12:28-14:30]: Beziehungsweise. Das was ich bisher wahrgenommen habe, wenn ich eben mit Menschen in unterschiedlichen Unternehmen gesprochen habe, meistens eher von technischer Seite als von strategischer Seite ist das, dass es auf der einen Seite in jedem Unternehmen Leute gibt die sozusagen die Bewahrer-Funktion übernehmen, das haben wir immer schon so gemacht. Und jegliches was man eben auch an Technik austauscht ist natürlich eine Fehlerquelle. Und da hat man sozusagen einen inneren Widerstand, der überwunden werden muss. Dann gibt es auf der anderen Seite Leute die vielleicht sagen, OK, die Technik ist neu und interessant, und wir wollen da Vorreiter mit sein. Die versuchen das Ganze dann also ein bisschen zu treiben, und dann prallen diese Fraktionen natürlich irgendwo ein bisschen auf einander. Das kostet auch immer wieder Geld. Und dann braucht man eben einen guten Anwendungsfall mit dem man das einmal prototypisch zeigen kann, dass die Vorteile wirklich ziehen. Und dann wird das eben weiter ausgebaut. Das ist eigentlich das, was man so hört. Die Frage ist: Selbst wenn man das jetzt hat, es muss ja auch eingesetzt werden. Es muss also jemand die Motivation wirklich haben da jetzt einen Nutzen draus zu ziehen. Und ich glaube da tun sich viele Unternehmen noch schwer, weil die eben ihr Marktfeld erschlossen haben, und ihre Gewinne oder was auch immer sie daraus, Umsatzziele haben, natürlich fließt da nachher Gewinn raus, die erfüllen im Moment und die jetzt die Befürchtung haben OK, wenn ich jetzt mehr in Technologie investiere, das muss ich nachher erstmal irgendwie wieder rein holen. Mit welchem Benefit gehe ich da raus, dass ich wirklich nachher einen Wettbewerbsvorteil

habe. Und das kann ein Techniker alleine sehr schlecht sehen, weil der hat von der eigentlichen Domäne, was die jetzt wirklich daraus holen können, nicht so viel Ahnung. Und jemand der in der Domäne arbeitet, weiß dann auch nicht, was bringt mir das jetzt wenn das doppelt so schnell ist. Also was kann ich damit machen. Ja, es ist doppelt so schnell, das ist kein Wert an sich.

R [14:31-15:00] OK. Und habt ihr da jetzt in der Zeit wo ihr euch mit dem Thema beschäftigt habt irgendwelche Veränderungen selbst schon beobachtet, also dass zum Beispiel die Unternehmen offener dem gegenüber werden, oder dass das doch einfach sehr langsam ist weil die sehr viel diese resistance to change haben, und das nicht einsehen wollen?

E7 [15:01-15:10] Also wenn man den Verkaufszahlen, die man hört, glauben darf, dann sind es ja immer mehr Unternehmen, die im Prinzip in die Richtung wechseln.

E6 [15:10-15:55]: Zumindest mal ausprobieren. Die, die halt früher ausprobiert haben und dann auch inzwischen die ersten sinnvollen, oder nicht im Probierstadium befindlichen, Lösungen bauen. Und aus meinem persönlichen Bekanntenkreis kann ich nur sagen, dass zumindest zwei Unternehmen jetzt da, also dass ich von zwei Bekannten von mir aus den jeweiligen Unternehmen mal angesprochen wurde, also das Interesse scheint langsam so durch zu diffundieren. Was dann da natürlich mit passiert weiß ich auch nicht, aber zumindest wurde dann doch mal – weil die ja natürlich auch wissen dass wir ein Stück weit Informationen darüber haben – schient da zumindest Interesse zu kommen.

R [15:55-16:05]: Und als Art Negativ-Szenario: könnt ihr euch irgendein Szenario vorstellen, wonach die Nachfrage nach Echtzeit-Analysen sinken könnte?

E6 [16:06-16:25]: Fehleranfälligkeit. Wenn mein System irgendwann mal steht, und dass ist irgendwie einer neuen Technologie geschuldet, warum auch immer, dann war's das.

E7 [16:26-16:39]: Oder es ist einfach irgendwann halt so schnell, dass die Verbesserungen nur marginal wären, und dann Kosten-Nutzen im keinem gesundem Verhältnis mehr stehen.

E6 [16:40-16:48]: Hätten wir wieder eine philosophische Frage, ob eine Sättigung bei der Geschwindigkeit gibt.

R [16:49-17:57]: Da haben wir ja schon relativ viel. Die greifen natürlich auch alle so ein bisschen in einander über, und wir würden jetzt zum nächsten Block gehen, da geht es um die Kosten konkret. Das haben wir ja auch schon ein bisschen angesprochen grade. Und zwar: Was ich jetzt auch gehört hab von verschiedenen – ich habe zum Beispiel auch Anwender interviewt – und die haben tendenziell gesagt diese RIMDB sind für viele einfach noch zu teuer. Vor allem weil die eben auf Lizenzgebühren basieren, und gerade mittelständische Unternehmen, wenn die eine open-source DB nehmen können, dann machen die das halt. Könnt ihr euch da irgendeine Möglichkeit vorstellen, wie dieses Dilemma gelöst werden kann? Also zum Beispiel haben wir auch als eine große – ja, als eine Idee, dass open-source DB, auch relationale in-memory open-source DB, eine Lösung wären.

E7 [17:57-18:01]: Die gibt es ja auch.

E6 [18:01-19:17]: Wäre natürlich die kostengünstigere Variante. Also die Frage ist wahrscheinlich nicht gut für uns zu beantworten. Wir sind eher mit der technischen Seite beschäftigt. Letzten Endes, man will ja damit Geld verdienen, in diesem Fall ist es die SAP, und die müssen damit Geld verdienen, und die bieten dafür eine gewisse Sicherheit und Support an. Und ein Unternehmen, das groß genug ist und dem seine Daten heilig genug sind, die bezahlen das. Und wenn ich eben Mittelständler bin und mein back-up liegt bei mir zuhause im Schrank so nach dem Motto, ganz zur Sicherheit nochmal, irgendwie auf dem Firmenrechner, und da nochmal, und das mache ich alle Monat, oder mein Admin macht das, und wir brauchen jetzt in-memory-Fähigkeit, ja und die wollen jetzt nicht die Geldscheine auf den Tisch legen. Das teure ist sozusagen der ITler, oder die drei ITler, oder die fünf ITler, die die beschäftigen, und dann lassen die die halt mal rum probieren. Und wenn das noch nicht sauber läuft dann brauchen die da auch nicht ein Unternehmen dahinter, sondern die probieren erstmal, und ich glaube da kann man auch nicht so viel machen. Letzten Endes will man natürlich wieder Geld an denen verdienen.

E7 [19:17-19:42]: Das kommt halt ohne dieses Service-Paket. Die open-source Geschichten, das ist dagegen ein großer Punkt. Und teilweise sind die kommerziellen Produkte auch ausgereifter als open-source Anwendungen. Und bieten da vielleicht noch ein, zwei Features die vielleicht anders nicht implementiert sind. Aber klar, das muss man halt vergleichen. Und gucken.

R [19:42-20:00]: Und könntet ihr euch vorstellen, dass da sich auch noch eine große Entwicklung ergibt bei den Kosten, also dass die auch einfach günstiger werden? Zum Beispiel, was weiß ich, hardware Komponenten weiter günstiger werden, oder sowas in die Richtung?

E7 [20:01-20:08]: Also ich denke hardware tendenziell eher als software .Hardware-Preise sind ja glaube ich fallend.

E6 [20:08-20:10]: Ich würde dir jetzt ganz stark widersprechen.

E7 [20:10-20:12]: Gut, kannst du machen.

E6 [20:12-21:32]: Also soweit ich weiß, sind da... also das teure letzten Endes in Bezug auf die HANA, dass man die eben lizenziert und nach verfügbarem Hauptspeicher bezahlt, und damit hat man natürlich irgendwo eine Grenze wenn man jetzt datenintensive Anwendungen darauf fahren möchte, und das jetzt nach Hauptspeicher bezahlen muss. Dann muss man schon wieder rechnen: Kann ich da denn meine Daten rein tun. Hauptspeicher-mäßig sind das einmal kosten – wenn ich die einmal gekauft habe, hab ich sie. Aber wenn ich dafür dann jedes Jahr irgendwie bezahlen muss, nur weil ich dann mehr Daten drin habe, dann ist es natürlich eine interessante Frage, wenn ich dann mal einen Strich ziehe und eine Summe drunter schreibe und mir dann die open-source Lösungen dagegen angucke. Und da ist wahrscheinlich irgendwo so einen sweet-spot zu treffen, ich glaube nicht, also natürlich die einmal-Anschaffungskosten sind sozusagen ein Hinderungsgrund es einfach mal auszuprobieren. Aber die Lizenzkosten auf Dauer sind natürlich ein Hintergrund das ganz strategisch zu betrachten wenn ich zu viele Daten habe. Gerade wenn man eben noch nicht weiß, ob man es denn dauerhaft machen möchte. Also da kommen dann glaub ich beide Hinderungsgründe zusammen: Zum Ausprobieren muss ich die hardware kaufen, also wenn ich dann den Erfolgsfall habe, dann muss ich aber die teuren Lizenzkosten bezahlen. Dann weiß ich aber auch noch nicht, nachdem ich die hardware gekauft habe, ob mein Anwendungsfall darauf läuft.

E7 [21:32-21:36]: Wobei, man nicht zwingend hardware kaufen muss, weil man auch Cloud-Lösungen...

E6 [21:37-21:47]: Ne, also das kann man ja erstmal probieren, aber da ist wieder das Problem: Welches Unternehmen steckt denn seine kritischen Daten wirklich irgendwo in die Cloud, woanders hin

E7 [21:47-21:48]: Ja klar, das ist eine andere Frage.

E6 [21:49-22:02]: Also ich glaub dass eben viele Einflussfaktoren, die da Kopfschmerzen bereiten. Und das führt dann auch zu einem gewissen Hemmnis es zu machen.

R [22:03-22:10]: Du meinstest jetzt die hardware wird an sich günstiger eher?

E7 [22:11-22:18]: Also die Preise für Hauptspeicher sind ja in den letzten Jahren gefallen wenn ich mich nicht täusche.

E6 [22:19-22:45]: Also ich kann dir nur sagen am Consumer-Markt war das nicht so. Ich hab vor drei Jahren gekauft, und vor einem Jahr gekauft. Und es war vor einem Jahr teurer als vor drei Jahren. Also das kommt immer drauf an wie sich der Weltmarkt für D-RAM entwickelt. Im Moment stagniert das eben, aber tendenziell wird es noch weiter wieder billiger werden. Also das ist auch einfach eine Markt [inaudible].

R [22:46-22:59]: Und könnt ihr mir ein Beispiel sagen für solche open-source Lösungen?

E7 [23:00-23:10]: VoltDB, ist so eine, oder? War Cassandra auch in-memory? Schreib mal Cassandra vielleicht nicht auf.

E6 [23:17-23:18]: Cassandra ist h-based, oder? Das ist keine relationale Datenbank.

E7 [23:18-23:26]: Ja eben, darum. C-Store. Und was ist denn das, Vertica.

E6 [23:26-24:54]: Also es gibt diverse, und auch, je nachdem was die halt jeweils dann können sollen oder nicht können sollen. Es gibt auch für MySQL einen Engine die auf Main Memory quasi laufen kann, die man da einbauen kann. Je nachdem, was man dann eben für Features brauch. Aber es wird dann eben auch hakelig. Man muss dann wissen was man tut, wie die sich verhalten, die Verzahnung... Also ich glaube der Vorteil, den sozusagen SAP mit HANA da liefern kann, und den sie auch liefern werden ist die Verzahnung mit der

Applikationsseite. Das wird natürlich dann irgendwann für Kunden von SAP das sein, was es dann interessant macht gegenüber einer open-source Lösung, wo das nicht garantiert wird. Ich hab mir noch ein paar angeguckt. Es gibt da noch Oracle x10, Konkurrenzprodukt. Microsoft hat auch eine, die hab ich mir auch mal ein bisschen angeguckt. Es gibt Diverse. Die sind aber auch nicht open-source.

E7 [24:54-24:55]: MonetDB

R [24:56-25:29]: Könntet ihr euch vorstellen, dass HANA, also das SAP oder die anderen großen Anbieter, Oracle oder Microsoft, selbst zu so einer open-source Lösung umschwingen und dann sich eher auf die Applikationsseite konzentrieren? Weil das könnte man sich ja auch als Business Model vorstellen. Oder wäre das für die...

E7 [25:29-25:45]: Im Bereich DB sehe ich's jetzt... Zeitraum 2025, richtig? Ne, sehe ich es nicht bis 2025. Dass SAP das DB-Geschäft aufgibt.

E6 [25:45-25:46]: Aufgibt?

E7 [25:46-25:50]: Ja also umschwenken zu open-source ist für mich...

E6 [25:50-25:53]: Also die Frage, wie lautet die Fragestellung genau?

R [25:59-26:14]: Ob zum Beispiel SAP oder andere große Anbieter von RIMDB selbst anfangen ihre open-source Lösungen, also ihre DB als open-source Lösungen zur Verfügung zu stellen, und dann praktisch als Business Model sich eher um die außen-rum Applikationen...

E7 [26:14-26:20]: Ach so, also ich sehe da keine große Entwicklung zumindest in die Richtung.

E6 [26:20-26:43]: Ich sehe es auch nicht. Also es gibt das Geschäftsmodell natürlich, was für kleinere Firmen funktioniert durchaus, seine eigentliche Kern-Software zu öffnen und kostenfrei zur Verfügung zu stellen. Und dann über customizing und support Geld zu verdienen. Einfach um den Enterprise-Support zu haben. Das funktioniert aber nur solange die Firma relativ klein ist.

E7 [26:44-27:09]: Also ich sehe einfach nicht den Bedarf dazu das kommerzielle Eigenprodukt jetzt zu open-sourcen. Wobei es natürlich bei beispielsweise bei VoltDB, das es dann eine kommerzielle Version gibt, und eine open-source für die community. Aber jetzt bei den SAP, Oracle, die du genannt hast, da sehe ich das jetzt nicht.

E6 [27:10-27:31]: Die Frage ist jetzt sozusagen, für Wissenschaft und Forschung gibt es dann immer Individual-Verträge, die so etwas ermöglichen, aber ein generelles open-source sehe ich jetzt nicht.

E7 [27:31-27:53]: Wobei jetzt SAP zum Beispiel ja auch vor ein paar Monaten, wenigen Jahren das JavaScript-Framework veröffentlicht hat auf Github, also genutzt wurde, wird, solchen Anwendungen [inaudible] ihre DB zu bauen.

E6 [27:54-27:59]: Wobei das ja auch nur quasi, nur in Anführungszeichen, ein Zusatzbaustein ist.

E7 [27:59-28:00]: Ja, natürlich. [inaudible]

E6 [28:00-28:11]: Also du hast nicht so unglaublich viel davon, wenn du nicht HANA drunter nimmst, also du würdest ein anderes... andere Bausteine.

E7 [28:11-28:14]: Ist nicht Teil der DB sozusagen.

R [28:15-28:22]: Aber es gibt auf jeden Fall zum Teil solche Initiativen, dass sie vereinzelt Sachen frei geben.

E6 [28:23-28:46]: Natürlich. Es gibt also... SAP ist auch ein großer Softwarehersteller, da arbeiten viele Menschen und viele Softwareentwickler haben da auch Beziehungen zur open-source Szene, ich meine man arbeitet mit open-source Tools, dann hat man natürlich auch ein Interesse daran, dass das was man geschrieben hat, und worauf man stolz ist von möglichst vielen verwendet wird, und wenn da eh kein Geschäftsmodell dahinter steht wie bei einer Zusatzbibliothek, dann wird das auch möglich.

E7 [28:47-29:07]: Ich mein man kann ja da auch profitieren. Ist ja eine win-win Situation. Wenn im Prinzip die community vielleicht Fehler findet, die man selber nicht gefunden hat, eventuell Dinge beiträgt, die man dann selber intern auch einbauen kann, die man auch so nutzen kann, das ist ja keine einseitiges Vorteilsverhältnis.

R [29:08-29:33]: Und jetzt mal so als ganz Negativ-Szenario für die Softwarehersteller – könnte es sein, oder wie sähe das aus, die Entwicklung dahin, wenn tatsächlich die open-source Lösungen die herkömmlichen Datenbanken verdrängen? Würdet ihr sagen, dass ist nicht möglich weil doch viele den Service so sehr schätzen?

E7 [29:33-30:06]: Wenn es so wäre, dann würde wahrscheinlich mit anderen Dingen halt eher Geld verdient werden, wie Anwendungen, Infrastruktur, Daten, Schulungen, Consulting Services, keine Ahnung. Würde der Fokus wahrscheinlich stärker auf diesen Dingen liegen. Oder halt Dinge die man eventuell neben der open-source DB braucht wie vielleicht ein streaming-System oder ein Hadoop-Cluster, um günstige, noch günstiger Dinge abzuspeichern.

E6 [30:06-31:36]: Ich seh's auch nicht wirklich kommen, weil es gibt open-source DB, die halt, die eine eins-A DB sind, aber keine Hauptspeicher-Spaltenorientierte DB. Das heißt, eine open-source Software, die halt da ist, und die auch ausgereift ist, und wo die von den Kinderkrankheiten befreit ist, und eine gute Community hat, die setzt sozusagen den Mindeststandard, den alle kommerziellen Anbieter mindestens in einem Punkt schlagen müssen, oder halt eine Niche finden, wo sie dann wirklich Geld für ihr Produkt nehmen können. Deswegen ist sozusagen eine open-source Lösung im Markt eigentlich für den Gesamtmarkt immer gut, weil sie die Messlatte sozusagen hoch setzt, aber da wo sich eben noch ein Markt entwickelt, und da sehe ich eben Hauptspeicher-DB, da ist erstmal noch genügend Zeit bis zu sagen... Potenzial, dass da etwas neu entwickelt wird, und dass man sozusagen die Lösung auch immer verkaufen kann. Und erst wenn das sozusagen alles ein alter Hut ist, dann wird es vielleicht doch irgendwann eine ähnlich gute open-source Implementierung geben, die das alles kann, was man da eben benötigt, und dann muss es da halt etwas anderes geben. Aber ich glaube in der Richtung ist auch gar nicht die Gefahr, jetzt die... der Hauptmehrwert eben daraus gezogen werden kann, dass eben Anwendungen auch von SAP-Seite und DB von SAP-Seite nicht verzahnt wird, aber zumindest auf einander abgestimmt wird.

E7 [31:37-31:47]: Gibt es denn, weißt du ob es Großunternehmen gibt, die eine open-source DB nutzen, für ihre ERP-Daten oder so?

R [31:47-33:25]: Nein, wäre aber auf jeden Fall sehr interessant. Solche Daten sind aber sicher ziemlich schwierig zu finden. Müsste ich vielleicht dann alle einzeln durchgehen. Aber das wäre auf jeden Fall interessant zu wissen.

Dann haben wir da schon viel besprochen und können gleich zum nächsten gehen. Und da geht's um gesetzliche Regulierung. Das heißt Datenschutz, Datensicherheit, und in dem Rahmen auch Standardisierung und Normierung von... zum Beispiel, was der Sicherheitsstandard sein muss. Zum Beispiel die EU hat ja eine neue Datenschutz-Richtlinie eingeführt, die in den nächsten Jahren implementiert werden soll, und das ist natürlich relevant für große Firmen. Ich würde aber sagen, generell sieht man die Relevanz auf den ersten Blick eher bei Consumer-Applikationen. Klar, vom Nutzer müssen die Daten geschützt werden. Wenn man sich aber auf industrieller Seite, Geschäftsprozess-Seite, was können da zum Beispiel Restriktionen sein, die jetzt von Gesetzgeber-Seite auf die Firmen einwirken? Schlagwort Industrie 4.0. [inaudible]

E6 [33:25-33:31]: Mir ist die Frage immer noch unklar. Von wem sollen Restriktionen gesetzt werden, von der Politik?

R [33:31-33:32]: Genau.

E6 [33:33-34:19]: Mit welchem Ziel? Die Landesinteressen zu schützen? Datenschutz wirkt ja auf verschiedene Art und Weisen. Jedes Unternehmen möchte sich selbst und seine Datenhoheit schützen. Es gibt Länder, die möchten gerne möglichst viele Daten sehen, Stichwort Industriespionage. Und die profitieren davon wenn man in Daten rein gucken kann. Und dann gibt es auch eben Länder, gut, dann... Wir müssen halt unsere Daten möglichst schützen, und in dem Spannungsfeld entscheidet sich das Ganze. Das ist auch mit politisch begründet wer da auch welche Daten sieht, oder nicht sieht, oder sehen möchte. Und das ist glaub ich ein Feld, was wir nicht wirklich beantworten können. Ich habe eine Meinung dazu...

E7 [34:20-34:34]: Also ich sehe da jetzt nicht dass durch neue Trends wie Industrie 4.0 viele neue Restriktionen auf uns zukommen ehrlich gesagt. Klar, bestehende Gesetze müssen angewendet werden.

E6 [34:34-35:23]: Also ich glaube die Politik ist bei dem was quasi informiertes Handeln hinsichtlich Datenhoheit angeht sowieso faktisch etwas hinterher. Und in der Hinsicht wird da irgendwas festgelegt, und irgendwas gemacht, das sieht man ja alleine am Safe Harbour Abkommen. Da werden dann Gesetze gekippt und es gibt keinen Nachfolger, und letzten Endes interessiert es überhaupt keinen, weil das Leben muss halt weiter gehen. Gestern oder vorgestern ist es ausgelaufen. Prinzipiell dürfen im Moment keine Daten in die USA übertragen werden. Ich bezweifle, dass sich irgendwo irgendwas geändert hat.

R [35:24-35:52]: Und was könnt ihr euch vorstellen... [inaudible]. In die andere Richtung jetzt wie von Regierungsseite sowohl die Anbieter als auch die Anwender von Big Data Solutions ganz allgemein tun können, damit man die Potenziale optimal nutzt?

E6 [35:52-36:02]: Sich raus halten. Die Politik kann sich da einfach raus halten. Ich glaube helfen können die nicht wirklich.

E7 [36:02-36:09]: Ja, man kann halt in Forschung auf diesem Gebiet investieren. Öffentliche Forschungsprojekte anstoßen.

E6 [36:09-36:14]: Ja, generell. Aber im Hinblick auf Datenschutz...

E7 [36:14-37:04]: Im Hinblick auf Datenschutz, also... Datenschutz ist ja... Also man kann vielleicht den Datenschutz, man muss halt einen Mittelweg finden, den Datenschutz so sicher zu gestalten, dass die Unternehmen darauf vertrauen vielleicht, dass ihre Daten sicher sind, wobei das sehe ich eher als Aufgabe bei den Unternehmen, und weniger bei der Regierung, auf der anderen Seite ist natürlich ganz wichtig, dass durch irgendwelche Gesetze oder Restriktionen da nichts behindert wird an Forschung oder gelähmt wird. Aber da ich die Datensicherheit, -schutz, in dem Bereich eher auf Unternehmensseite sehe ist da ein potenzieller Mehrwert gering.

R [37:04-37:50]: Dann zu einem ganz anderen Themenbereich, und zwar das Netzwerk. Das heißt, es ist ja um die ganzen aus dem IoT zu sammeln muss man die natürlich auch in irgendeiner Form übertragen. Gerade im öffentlichen Raum. Also in einer eigenen Fabrikhalle ist das natürlich einfacher, aber im öffentlichen Raum. Könntet ihr euch da ein Szenario vorstellen wonach sich zum Beispiel das IoT schneller ausbreitet als das Netzwerk mithalten kann, oder was da mögliche Lösungen wären, damit das Netzwerk mithalten kann praktisch? Also wenn man jetzt Berlin anschaut: OpenWiFi ist immer noch seit Jahren nicht verfügbar.

E7 [37:50-38:00]: Also da wo ich her komme habe ich auch keinen Handyempfang oder mobiles Internet, also da ist ein Funkloch.

E6 [38:00-38:03]: Bei mir in der Heimat gibt's da auch so diverse Fälle.

E7 [38:03-38:09]: Können wir denke ich noch viel machen. Ich glaube Deutschland ist ja da auch nicht auf den Spitzenplätzen was den Ausbau angeht.

E6 [38:09-39:00]: Was aber auch wieder an der Politik liegt, ich meine wer teilt denn... Die Netze gehören der Deutschen Telekom, die waren mal ein Staatsbetrieb, und man merkt das auch noch. Die haben halt das Netz, und die bauen es da aus wo die Konkurrenz halt da ist. Und da wo die Konkurrenz nicht da ist wird's eben nicht ausgebaut. Es gibt halt Konkurrenten am Markt die jetzt eigene Glasfaser-Leitungen ziehen, damit das Ganze ein bisschen unangenehmer machen für die bisherigen Marktteilnehmer. Das wird sich halt jetzt langsam weiter entwickeln. Also ich glaube der politische Anspruch flächendeckend – ich weiß jetzt gar nicht was die Kennzahl gerade ist, die flächendeckend rausgegeben wird – der wird nicht erreicht werden. Dieser Anspruch ist sicherlich hilfreich, weil nur die Politik unseren großen deutschen Netzbetreiber dazu zwingen kann Dinge zu tun.

E7 [39:00-39:20]: Das sind glaube ich hauptsächlich die dünnbesiedelten Gebiete, in denen die Bevölkerungszahl nur sinkt, also zum Beispiel da wo ich herkomme. Ich trage dazu aktiv bei. Kann ja nicht besser werden.

E6 [39:20-39:38]: Also industrielle Anwendungsfälle war halt die Frage dahinter, fallen mir jetzt keine direkt ein. Aber man sieht ja schon wenn die Leute einfach nur HD-Streams machen wollen ist ja meistens schon Ende auf dem Land.

E7 [39:39-39:44]: Also kleine, mittelständische Unternehmen, in Brandenburg zum Beispiel, haben auch das Problem, dass das Internet zu langsam ist.

E6 [39:45-40:43]: Aber haben sie wirklich ein Datenproblem? Also die haben das Problem, dass sie einfach keine gute Verbindung haben, aber woran fehlt es denen? Fehlt es denen an Volumen? [inaudible] Ich glaube nicht, also ich weiß es nicht. Der nominelle Datenstrom, den du als Unternehmen da durch schickst ist ja nicht irgendwie... Also wenn du jetzt wirklich an IoT in die Richtung denkst, hast du jetzt keine HD-Videos, sondern das sind irgendwelche Kommunikationsdaten, sonstiges, das ist ja eigentlich ein Witz an Datenmenge. Ich seh's nicht, dass das irgendwie ein Problem ist. Das Problem ist glaub ich eher nur die Infrastruktur bis an dieses Gerät und an die Kommunikationsschnittstellen zwischen den Geräten – nicht die Bandbreite. Erstmal nicht die Bandbreite!

R [40:44-41:13]: Und dann geht's noch zum Schluss als Basis von diesen ganzen IoT-Daten, und zwar diese Prognosen zum Anstieg der Sensoren, und diese explosionsartige Vermehrung von IoT. Haltet ihr daws eher für einen Hype, oder seht ihr da schon Tendenzen, dass das wirklich so, wie zum Beispiel Gartner das jetzt vorhersagen...

E7 [41:13-41:15]: Ist das jetzt gerade auf dem Peak?

R [41:15-41:19]: Ja genau, ganz oben.

E7 [41:20-42:40]: Ja also denke ich, gerade aktuell schon ein Schlagwort, Trend, das oft und gerne benutzt wird. Wird wahrscheinlich auf dem Gartner Hype Cycle wieder runter gehen entsprechend. Aber ich denk schon, dass das langfristig ein Thema wird. Weil halt dadurch es möglich ist noch genauer Dinge vorherzusagen, berechnen, analysieren, schneller zu reagieren auf sich ändernde Umstände, besser zu optimieren. Also das kann man schon in bestimmten Bereichen nutzen sicher, um da am Ende irgendwie einen Benefit raus zu ziehen, Kosten zu sparen, oder Umsatz zu steigern, oder was auch immer. Man hört ja jetzt zum Beispiel von Beispielen wo das dann entsprechend eingesetzt wird. Zum Beispiel Kosten sparen wenn da eine Fräse, die bricht immer ab, und die wird gekühlt durch Wasser-Strahle, und man sieht nicht wenn die abbricht. Aber jetzt wurde da ein Akustik-Sensor angebracht, und der nimmt das Abbrechen akustisch wahr, und so kann man sofort darauf reagieren, und stoppen, und einen neuen Schrauber oder Bohrer einsetzen ohne dass da Ausschuss produziert wird bei den nächsten fünf Teilen wo es noch nicht bemerkt wurde. So Szenarien gibt es halt.

E6 [42:40-43:20]: Ich bin da etwas skeptischer unterwegs. Meiner Meinung nach... Meine Frage dazu wäre, was hat das mit dem IoT zu tun? Überhaupt nix, denn da brauchst du ja eigentlich nur einen Sensor an der Maschine. Und der muss einen Alarm geben. Da muss nichts ans Internet, das ist alles lokal. Ich gehe davon aus – natürlich ist es eine Bewegungsrichtung, die langfristig Optimierungspotenzial bietet, gerade wenn Firmen untereinander geteilte Herstellungsprozesse haben, dass sie diese Daten besser austauschen können. Das ist aber eigentlich – ja, es ist ein Hype-Wort unter das viele andere Entwicklungen jetzt runter gepackt werden, weil sie sich dann besser verkaufen lassen – meiner Meinung nach.

E7 [43:21-43:25]: Wenn der Bohrer bricht, dann kannst du direkt den Roboter los schicken, der das Teil zu dir fährt.

E6 [43:26-43:47]: Ja, vielleicht. Das ist aber alles innerhalb der Firma. Also du müsstest jetzt sagen OK, dir bricht diese Säge weg, und das Gerät hat aber einen Tag vorher schon erkannt, dass es bald bricht und hat schon mal bestellt, das Teil was du nicht hast. Also wenn du es jetzt soweit spinnst, dann hast du halt irgendwann wirklich...

E7 [43:47-43:51]: Es gibt da ja auch verschiedene Definitionen von IoT, was ist das überhaupt, also da kann man...

E6 [43:51-44:24]: Weil das gerade hyped baut halt jeder, das ist genauso wie Big Data, wenn du jetzt irgendwas gerade gemacht hast wo halbwegs viele Daten drauf sind, dann schreibt jeder Big Data drauf. Jetzt ist es halt IoT. Und als nächstes wird es dann noch etwas anderes wieder. Also die Einzelentwicklungen sind sicherlich alle sinnvoll. Die meisten Leute die über solche Sachen reden ist halt politisch motiviert. Auch wenn du mir jetzt eigentlich widersprechen musst.

E7 [44:24-44:25]: Nein nein.

R [44:25-44:35]: Und in welchen Bereichen seht ihr die größte Relevanz von solchen Technologien, spezifisch auf die Sensoren bezogen?

E7 [44:25-44:45]: Sicherlich bei hochpreisigen Produkten wird das sicherlich eine Rolle spielen, weil da halt großes Einsparpotenzial denk ich ist.

E6 [44:46-48:06]: Also ich denke halt immer, die Menschen sind ja jetzt auch nicht blöd. Da wo es pain points gibt, wo irgendetwas im Argen ist, da arbeiten die jetzt auch schon dran, auch ohne IoT und sonstigen Dingen. Und wenn jetzt ein Fertigungsprozess ineffizient ist weil irgendwelche Dinge kaputt gehen, dann werden sich die Leute da eine Lösung für überlegen. Und klar, vielleicht gibt es einfach sozusagen im Moment irgendwo Stellen wo man jetzt sagt OK, hier ist jetzt Ende, hier kann ich als Einzelperson, oder als kleine Firma nichts mehr dran machen, weil ich kriege diese Daten jetzt nicht an die Maschine dran; die sind einfach inkompatibel untereinander, und als eine Firma kann man da nicht den Hersteller sozusagen sagen, macht das jetzt mal kompatibel. Also das wird sich denke ich verbessern, gerade im industriellen Sektor, dass es offenere Schnittstellen im weitesten Sinne gibt, dass Maschinen untereinander, miteinander reden können. Das wird denke ich passieren, ich würde sagen der industrielle Sektor sind ja auch hochpreisige Dinge.

Ich sehe es nicht so sehr im Consumer-Bereich. Ich sehe auch noch nicht, dass mein Kühlschrank automatisch mein Essen nach bestellt, weil warum sollte ich das. Vielleicht möchte ich mir ja lieber was anderes selbst im Supermarkt aussuchen.

Aber zum Beispiel die Logistik-Branche hat sich ja massiv, auch wenn das jetzt wieder nicht IoT genannt wurde, die hat sich ja massiv durch IT professionalisiert. Dass ich wirklich sehe: Wo ist mein Paket, wo geht es hin, ich kann noch anrufen während das Paket unterwegs ist und es woanders hin schicken. Oder über eine Webseite sagen: Bitte bei meinem Wunsch-Nachbarn abgeben, oder doch nicht. Da habe ich als Privatmensch einen Komfortgewinn, da hat das Unternehmen eine bessere, durchs Tracking, bessere Verfolgbarkeit. Gerade auch für Industriegüter die sehen können: OK, dann und dann kommt es. Also ich denke: Solche Prozessoptimierungen werden dann nicht nur in der Logistik stattfinden, sondern in allen Bereichen. Da glaube ich passiert dann einiges.

Da kriegt der Consumer vielleicht auch was von ab, wenn es da irgendwie... weiß ich nicht, ich fände es zum Beispiel gut wenn ich Fleisch kaufe, und dann genau sehen könnte: Das ist an dem Tag von da nach da gefahren; OK, muss man den Daten auch wieder vertrauen. Kommt von dem und dem Hof, so sah die Kuh aus nach dem Motto. Warum nicht? Das wäre technisch alles möglich! Muss man natürlich auch irgendwo bezahlen, aber man könnte natürlich da mit Technologie sicherlich was machen. Also so sadistisch oder komisch sich das jetzt anhört, aber die Kuh ist auch am Ende auf der Farm, oder auf dem Bauernhof, dann fährt die zum Schlachter, also wenn man es wirklich auf die Spitze treiben würde könnte man davon ein Foto machen. Und das mit dem QR-Code ins Internet laden, und wenn ich die Fleischbrocken darauf halte kann ich mir angucken: Die Kuh wurde zu dem Schlachter gefahren, und wurde um die Uhrzeit um die Ecke gebracht, und dann von dem zerlegt... Also jetzt mal übertrieben dargestellt. Aber ich könnte diesen Weg verfolgen.

Vielleicht interessiert es mich weniger beim Essen als bei meiner Kleidung, oder weiß ich nicht was. Natürlich macht es das teuer, aber ich glaube der Markt wird zeigen wo da auch ein Bedarf für da ist. Also beim Essen würde es mich zumindest interessieren: Wo kommt's her?

R [48:06-48:12]: Und gerade sind ja auch parallel Trends regional, FairTrade...

E7 [48:13-48:28]: Nebenbei der Kollege hat sich einen Luftfeuchtigkeitssensor gebaut und neben das Fenster gehängt, und der kriegt jetzt gesagt wann er lüften muss. Das gibt es auch schon produktiv.

E6 [48:28-48:46]: Oder du machst es einfach jeden Tag. Ne, aber ich verstehe schon. Aber es wäre zum Beispiel dann schön, wenn mein Haus dann anfängt selbst zu lüften, ich kann das ja nur machen wenn ich zuhause bin im Moment.

R [48:46-48:48]: Genau, also wenn du im Urlaub bist zum Beispiel...

E6 [48:48-48:54]: Dann wird auch nicht gelüftet. Aber ich will ja auch nicht, dass es dauernd lüftet, dann kommt der Einbrecher und lüftet dauerhaft.

E7 [48:53-49:00]: Das Haus kann doch lüften wenn du im Urlaub bist, wenn es nur die Fenster kippt?

R [49:00-49:03]: Wahrscheinlich ist das versicherungstechnisch schwierig.

E7 [49:03-49:07]: Wenn es kaputt ist geht es die ganze Zeit auf und zu.

R [49:08-49:36]: Und als Abschlussfrage: Gibt's irgendwelche Überraschungen, die ihr euch vorstellen könnt, was vielleicht nicht ganz so viel besprochen wird, aber wo ihr mal drauf gestoßen seid, die sich im Bereich Big Data und IoT innerhalb der nächsten zehn Jahre entwickeln könnten? Vielleicht auch irgend ein Thema, das ihr besonders wichtig findet, aber gar nicht so bekannt ist?

E7 [49:36-49:48]: Was man zuletzt ja immer, ab und zu sieht, sind diese großen Brillen, dieses augmented reality, oder virtual reality. Also diese großen...

E6 [49:49-49:51]: Beides! Oculus Rift, oder HTC-Vive, oder...

R [49:51-49:53]: ...Microsoft Hololens...

E7 [49:53-50:15]: Genau, genau. Ich bin gespannt, was damit am Ende passiert. Ob das irgendwie, weiß ich nicht, ob das benutzt wird vielleicht außerhalb von Computer-Spielen. Aber da hab ich, also kann ich schlecht einschätzen, habe ich nicht länger drüber nachgedacht, aber finde ich ganz spannend wohin das geht. Hab so ein Ding noch nie aufgehabt.

[50:15-50:31]: [brief exkursus conversation on augmented/virtual reality glasses]

E7 [50:32-50:45]: Das könnte eine Überraschung werden wenn das wirklich in Unternehmen eingesetzt wird. Zumindest für mich, das ist natürlich subjektiv was eine Überraschung ist.

E6 [50:45-52:17]: Zehn Jahre ist eine lange Zeit. Also was ich mir vorstellen könnte, dass irgendwann relativ viele Daten gesammelt werden, und dann kriegt man irgendwann ziemlich interessante Fakten über die Bevölkerung raus, die dann auf einmal aufschreit, dass sowas überhaupt möglich ist mit diesen Daten. Das könnte ich mir vorstellen. Und dann wird der Kampf Datenschutz und ach, ist doch alles nicht so schlimm, nochmal zelebriert. Und dann kann jeder sich selber überlegen auf welcher Seite er da steht. Also es wird nichts völlig Brisantes sein. Aber ich glaube wenn man zum Beispiel... Ich mein, Daten sind auch jetzt schon vorhanden. Man kann in Verkaufsdaten von Supermärkten jede Menge interessante Dinge ablesen. Die sind zum Teil anonymisiert, zum Teil natürlich auch nicht. Wenn ich jetzt mit meiner Deutschlandcard da einkaufen gehe, die Märkte wissen schon, was sie da machen. Die gehen ja auch auf gewisse Art und Weise sensibel damit um. Sie strukturieren halt nur nach innen so, dass sie ihren Verkauf optimieren, und treiben kein Schindluder mit den Daten, weil sonst geht der Schuss natürlich gewaltig nach hinten los. Und wenn man das jetzt auf alle Bereiche überträgt, zum Beispiel Verkehrsüberwachung komplett damit macht, den man ja eigentlich nicht überwachen möchte, aber zur besseren Steuerung an den Ampeln irgendwas erfasst und sonstige Dinge macht, und dann vielleicht irgendwann Bewegungsprofile auftauchen. Ich glaube da können schon noch entweder positive oder negative Überraschungen bei rum kommen. Wenn erstmal genügend Daten irgendwie da sind.

R [52:17-52:21]: ...und wenn man dann auch weiß, wie man die dann vielleicht analysieren kann.

E6 [52:21-52:25]: Das kriegt man ja schon hin.

E7 [52:25-52:28]: Mit relationalen in-memory DB.

R [52:29-52:35]: Super, dann haben wir es schon, vielen Dank!

Transcript E8

Date: 08.02.2016

Duration: 50:11 [mm:ss]

Modus: telephone interview

Interviewee:

Expert 8 [E8]: Professor for Database Systems and related topics at a German Technical University of Applied Sciences; member of research and expert groups on big data analytics.

Interviewers:

Julia Reichelt, Researcher [R]

[00:00-05:00]: [introductory remarks]

R [05:00-05:38]: Dann können wir anfangen. In der ersten Gruppe geht's um die Nachfrage von Echtzeit Reaktionen, also das ist der erste Einflussfaktor. Nachfrage von Echtzeit-Reaktionen, Echtzeit-Analysen, und ganz allgemein die Beschleunigung von Services. Da wäre jetzt meine erste Frage wer Ihrer Meinung nach überhaupt sowas braucht? Wer macht Echtzeit-Analyse? Welche Arten von Firmen, und innerhalb der Firma welche units, können sowas wirklich gebrauchen?

E8 [05:39-05:43]: Echtzeit jetzt im Bereich welcher Latenz?

R [05:44-05:48]: Das ist Ihnen überlassen.

E8 [05:49-05:56]: Das ist aber schwierig.

R [05:57-06:09]: Ich habe es nicht genauer eingegrenzt. Ich habe jetzt auch schon von mehreren... von vorherigen Interviewpartnern unterschiedliche Antworten bekommen im Bereich auf unterschiedliche Latenz.

E8 [06:09-09:50]: Genau, also Sie haben natürlich... Was können Hauptspeicherdatenbanken [HSDB] gut? Sie können also sowohl transaktional als auch analytische Daten in kurzer Latenz verarbeiten. Und dadurch haben Sie... Wo hat man denn jetzt diese sehr kurzen Transaktionszeiten?

Und da gibt's halt... das typische Szenario ist im Telekommunikationsbereich wenn Sie also das Monitoren und das Managen von Handydaten, von Handymasten, auswerten wollen, dann brauchen Sie sehr kurze Latenz, weil Sie natürlich in dem Moment wo Sie sich einwählen während Sie Auto fahren auch das Gespräch nicht abbrechen darf. Und aber Überlastung natürlich auftreten können. Sie haben's natürlich im Börsenhandel. Da gibt's halt Speziallösungen von Oracle mit 30 Millionen Transaktionen pro Sekunde, die sehr teuer sind, aber wo natürlich das was jetzt allgemein HSDB genannt wird schon längst passiert.

Sie haben es in Routern, wenn der Router also lernen möchte welche Viren-Muster auftreten, dann haben Sie also auch hier transaktionale Daten, nämlich die Pakete die vorbei rennen, und der Classifier, der feststellt ob da ein Virus im IT-Paket drin ist oder nicht.

Das sind aber nicht die Ihnen jetzt die SAP verkauft. Die SAP verkauft Ihnen halt derzeit Szenarien, wo man sehr gut analysiert hat was sind die Workloads im ERP-Bereich, was sind die Workloads im CRM-Bereich, und hat festgestellt, dass die über 90% lesend sind selbst wenn wir hier schreibende Workloads anschauen. Und das sind halt so typische Szenarien, dass man halt interaktiv mit Menschen agiert. Dass man also an der Kasse sofort sozusagen die Empfehlung gibt: Kriegt der jetzt noch einen Bonus.

Wir schauen uns HSDB an vor dem Hintergrund, dass sie zwar in der Anzahl der Cores die pro Recheneinheit existieren beschränkt sind, dass sie aber sich trotzdem deswegen immer Lernverfahren eignen. Und man in maschinellen Lernverfahren natürlich sehr schnell Zulaufzeit realisiert. Und die Anwendung, die wir hier propagieren heißt text mining, das heißt also während Sie den Satz noch schreiben erkennt das System sofort worüber Sie schreiben. Ob Sie also über einen Stimmungszustand als kranker Patient eines Krankenhauses schreiben, ob Sie das Produkt nicht gemocht haben, ob Sie das Rock-Konzert von Eric Clapton nicht so toll fanden, und dann kann man natürlich schon während des Schreibens Feedback zurück geben, ob das, was die Maschine verstanden hat auch wirklich so war – die Maschine kann sich natürlich auch irren wenn sie das

interpretiert, was Sie im Text geschrieben haben. Oder Sie erhalten als Patient während des Schreibens korrelierende Daten hinzu über Ihre Vitaldaten beispielsweise, die gerade zu dem Zeitpunkt waren, über den Sie schrieben. Also: Ich hatte gestern Abend Kopfschmerzen. Dann sehen Sie sofort sozusagen wie Ihr Blutzucker, Gewicht, und so weiter gestern Abend war.

R [09:51-09:57]: Und das ist dann praktisch auch ein Bereich wo Sie jetzt sagen würden der ist im Kommen?

E8 [09:57-10:58]: ...Menschliche Interaktion unter 300 Millisekunden. Ich mein, wenn Sie sich Google angucken, die Google-Architektur war sehr perfekt, und auch Bing, basiert nur auf HSDB. Und HSDB sind ein auslaufendes Modell. Das nächste Modell was wir bereits jetzt schon am Markt finden, aber keiner bezahlen kann, ist Non-Volatile-RAM, also Non-Volatile Main Memory. Also mein Memory wird selbst wenn Sie jetzt da den Stecker ziehen trotzdem da erhalten bleiben. Und das ist momentan noch ein bisschen teuer. Und die Frage ist ja eigentlich eher – was machen wir, wenn's NV-RAM gibt? Also die Frage, die Sie eigentlich stellen müssen, ist was sind Anwendungsszenarien für NV-RAM, weil die Szenarien für HSDB sind die, die Sie im SAP-Kontext sehen, oder EXASOL-Kontext, oder aber es sind die, die Sie im Google-Kontext sehen, oder die Sie explorativ sehen. Die ganze Google-Architektur basiert auf HSDB die Batterie-gepuffert sind. Und deswegen beim Ausfall auch hier erhalten bleiben.

R [10:59-11:39]: Genau solche Sachen sind für mich sehr interessant, weil ich eben auch angucken will: Wie entwickeln sich Konkurrenztechnologien? Es geht also nicht nur darum zu gucken genau was machen die in-memory Datenbanken (IMDB), also was könnten auch andere Technologien sein, die vielleicht den Markt übernehmen?

Als Zeithorizont, das habe ich leider vorhin vergessen zu sagen, habe ich das Jahr 2025 angesetzt. Wie würden Sie sagen, sieht das im Jahr 2025 aus? Nehmen Sie an, dass sich der NV-RAM sich dann schon durchgesetzt hat?

E8 [11:39-12:01]: NV-RAM für Non-Volatile. Keine Ahnung, 2025, jetzt haben wir 2015, das heißt also wir hätten 1990 erraten müssen wie die Welt 2015 aussieht – keine Antwort. Gebe ich nicht. Das ist unseriös. Gucken Sie sich Matrix an. Punkt.

R [12:01-12:17]: Ja, also das ist genau der Kern dieser Methode – es geht nicht darum eine exakte Vorhersage zu machen, sondern wir wollen einfach den Möglichkeitshorizont aufspannen. Und gucken was vielleicht möglich und vorstellbar wäre. Deswegen...

E8 [12:17-12:46]: ...Ja, ja, ja, aber ich meine da können Sie auch einen Zehnjährigen fragen. Überlegen Sie mal, 25 Jahre, das ist wie wenn Sie 1990, da habe ich auf einem C64 programmiert. Hätten Sie jetzt vorhersagen wie unsere Welt jetzt ist. Also 1990 gab es Zurück in die Zukunft, 4, oder 3, oder so. Da flog der genau ins Jahr 2015 oder 2016. Und das macht ein Film in Hollywood, und auch bei Zurück in die Zukunft waren einige Vorhersagen falsch. Also die Hover-Boards zum Beispiel gibt's nicht.

R [12:46-12:51]: Ja, es geht auch nicht darum, dass es wirklich 100% stimmt, sondern wir wollen einfach...

E8 [12:49-12:56]: Ne, ne, das mache ich nicht. Sorry, mache ich nicht. Ist unseriös.

R [12:57-13:46]: Gut, dann können wir mal versuchen weiter zu machen, aber ich weiß nicht ob wir dann noch so viel hier heraus holen können, weil genau darum geht's bei der Szenarioanalyse.

Dann können wir gleich zum nächsten Fragenblock gehen, das sind die Kosten. Aktuell laufen die meisten IMDB ja über Lizenzsysteme. Können Sie sich da auch ein anderes Modell vorstellen, dass vielleicht innerhalb der nächsten Zeit eben sich entwickeln könnte? Das Problem ist ja, dass die Lizenzsysteme zum Beispiel für mittelständische Unternehmen sehr teuer sind. Deswegen können die das nicht benutzen.

E8 [13:46-14:04]: Ja, aber das stimmt ja nicht. Die meisten... Der Markt für HSDB... Wer sind denn eigentlich in Ihren Augen, anders gefragt, frage ich erstmal Sie: Wer sind denn in Ihren Augen die größten Anbieter von HSDB?

R [14:04-14:10]: Ich würde sagen ich spezialisiere mich auf keine.

E8 [14:10-14:12]: Welche kennen Sie?

R [14:13-14:18]: Ich kenne Oracle, IBM, SAP, und es gibt noch ein paar mehr.

E8 [14:18-14:21]: Und wie heißt das Produkt bei IBM?

R [14:21-14:24]: Das weiß ich jetzt spontan nicht. Ich bin kein Techniker und darum geht es bei mir auch nicht.

E8 [14:24-14:32]: Aber das sollten Sie schon kennen, wenn Sie darüber schreiben. SAP HANA ist einfach. Wie wird denn SAP HANA vertrieben?

R [14:33-14:36]: Das weiß ich nicht und darum geht es bei mir gar nicht...

E8 [14:35-14:39]: Das war doch genau Ihre Frage, die Sie an mich stellen.

R [14:40-14:44]: Es geht um die Lizenzsysteme...

E8 [14:43-15:12]: Und SAP HANA wird lizenztechnisch sowohl als eine Lizenz vertrieben, die pro 65GB Hauptspeicher bezahlt werden muss, der genutzt wird, oder zum Beispiel über Cloud-Services, oder sogar versteckt, nicht Cloud-Services, sondern Services, die da oben drauf sind. Genauso bei IBM. Und das ist keine Zukunftsmusik mehr, sondern das ist einfach Marktnotwendigkeit wie Sie es gerade beschrieben haben.

R [15:12-15:15]: Und was wären vielleicht andere Lösungen?

E8 [15:15-15:19]: Wie bezahlen Sie gerade bei Google?

R [15:19-15:21]: Mit meinen Daten.

E8 [15:21-15:37]: Ja, mit Daten ist das eine und das andere? Daten ist schon mal gut, aber es gibt noch was anderes bei Google. Oder andersrum, wie bezahlen Sie bei Facebook? Nicht nur mit Daten, sondern mit...

R [15:37-15:41]: Naja, indirekt über Werbung, die sie schalten.

E8 [15:41-16:02]: Genau, und was ist das dann, wenn Sie dann irgendwie fünf Minuten diese Werbung angucken durch Zufall? Lebenszeit. Sie bezahlen mit Lebenszeit. So. Und Daten ist eigentlich zu schwach, sondern was geben Sie in dem Moment Facebook oder Google von sich?

R [16:03-16:08]: Tja. Informationen über Verhalten.

E8 [16:08-16:37]: Ja. Ein Modell von sich selber. Und genau so ein Modell werden auch Firmen eingehen müssen, sie werden es auf gar keinen wollen, aber wenn sie... Firmen werden sich es sehr genau überlegen, weil sie werden's wahrscheinlich dann erst finanzieren. Aber diese Modelle entstehen. Und Daten-gegen-Daten-Modelle, gibt's schon seit Ewigkeiten. Das macht schon der GfK, oder Ebay, die machen das seit Ewigkeiten. Lebenszeitmodelle wird's auch geben.

R [16:38-16:52]: Da verstehe ich jetzt aber nicht genau im Bereich von in-memory Datenbanken - würden Sie dann sagen die Kunden müssen ihre Daten dafür frei geben, dass sie diese Datenbanken nutzen können?

E8 [15:53-17:25]: Sie haben derzeit Google Drive. Und Google Drive ist nichts anderes eine in-memory Datenbank für Web-Anwendungen, wobei sie halt hier einen key-value-store haben. Sie haben nicht das [inaudible] Datenmodell, sondern ein key-value-Datenmodell. Aber de facto übergeben Sie hier alle möglichen Daten an Google. Und bei Drive kriegen Sie noch nicht mal Werbung. Aber Drive lernt aus den Korrelationen der Wörter bessere Modelle um Ihnen in einem anderen Produkt, nämlich der Google Suche, bessere Werbung einzuschalten.

R [17:26-17:51]: Gut. Dann können wir schon zum nächsten gehen. Für wie relevant halten Sie open-source Datenbanken, also relationale in-memory Datenbanken, im Vergleich zu den kommerziellen Anbietern.

E8 [17:52-17:55]: Welche open-source Datenbank im Bereich Hauptspeicher kennen Sie denn?

R [17:55-18:06]: VoltDB wurde mir erzählt zum Beispiel.

E8 [18:06-18:23]: Volt? Achja, Volt, OK. Das ist aber nicht relational. OK. Nicht relational. Das ist die einzige die Sie bisher kennen?

R [18:23-18:33]: MySQL hat soweit mir gesagt wurde auch eine... Könnte man eine Hauptspeicher drauf setzen.

E8 [18:30-20:40]: Ist keine Hauptspeicher. Ne ne, das ist Quatsch, ne. Das ist sozusagen... Das ist Quatsch bei MySQL, da haben Sie ein Hauptspeicher-Modul, und wenn Sie einen Stecker ziehen, dann ist es weg. Die... Das was HANA macht basiert auf... Erstmal: Warum ist VoltDB nicht so wirklich relational? Die ist natürlich schon relational, die ist von Mike Stonebraker entwickelt worden und so weiter. Aber, großes Problem bei VoltDB ist, dass sie natürlich mit kontinuierlichen snapshots arbeitet. Und das hat ein bisschen was mit Transaktionsverhalten zu tun. Also ist nicht so transaktionsfähig wie vielleicht eine SAP HANA. Also Sie können da einfach nicht so viele Transaktionen drauf jagen. Insofern ist sie schon ACID-kompatibel, sie ist auch schon relational, aber natürlich wird sie nicht das leisten was jetzt andere Datenbanken leisten.

Und was sie natürlich haben ist MonetDB. Das ist das worauf in den... Ende der 90er hat das das Team in Amsterdam entwickelt. Und die Entwickler von HANA haben sich damals MonetDB sehr genau angeschaut, und haben es de facto zum großen Teil zunächst kopiert. Die Neuigkeit bei HANA war eigentliche her, dass HANA den Mut hatte in den Markt rein zu gehen. Dass Hasso Plattner sich 2006 auf einer Konferenz hinstellt und sagte: Hauptspeicher wird irgendwann mal so billig, dass wir hier auch mehrere Terabyte Hauptspeicher bei unseren Kunden einsetzen werden. Das hat damals nicht wirklich jemand geglaubt.

Und das andere was Sie haben ist die Firma EXASOL in Deutschland, die auch schon seit 15 Jahren auf dem Markt ist.

R [20:41-20:44]: Aber die sind nicht open-source...

E8 [20:43-20:49]: Und die ein sehr gutes Produkt... Es gibt ein sehr gutes Produkt von über 100 Kunden. Denen geht's ziemlich gut.

R [20:49-20:53]: Und ist das eine open-source Lösung? Das habe ich jetzt nicht so verstanden gehabt bisher.

E8 [20:54-20:58]: Das ist eine Community-Lösung derzeit. Open-source ist es nicht, aber es ist eine Community-Lösung.

R [20:59-21:29]: Innerhalb... Ich muss jetzt wieder die Zukunft eingehen: Haben Sie das Gefühl, diese offenen Lösung sage ich jetzt mal allgemeiner – vielleicht nicht open-source, aber zumindest Community-basiert – gewinnen an Relevanz? Oder, dass die niemals so ganz den Markt erobern können?

E8 [21:30-23:54]: Das ist eine Frage was man will nochmal. Diejenigen, die wirklich das Bedürfnis haben, dass sie hoch interaktiv in weniger als einigen 100 Sekunden Latenz arbeiten müssen – einigen 100 Millisekunden Latenz, Entschuldigung, haben meistens auch das Geld hier diese Leistung zu erbringen.

Und Sie haben in München halt das Hyper-System, was auch eine open-source Datenbank, die bestimmt technisch nicht schlecht ist, aber die sich überhaupt nicht durchsetzen konnte, weil es ein Uni-Projekt ist. Da ist einfach kein Support dahinter. Das wird von drei Entwicklern gebaut, oder vier, und dann gibt's noch zwei Profs die das machen. Insofern: Das ist eine schwierige [inaudible].

Bei MonetDB steckt eine Firma dahinter, bei EXASOL eine sehr große Firma sogar, also im Vergleich zu anderen kleinen Unternehmen. VoltDB ist halt so ein Ding von Mike Stonebraker und Sam Madden, was die also gemacht haben, wo man halt hofft, dass es die Community weiter treibt. Aber letzten Endes, wenn Sie ein Unternehmen sind brauchen Sie vollen Produkt-Support. Und dafür kann halt nicht nur ein Apache-Community dahinter stecken.

Ja, und auf der anderen Seite gibt es halt viele kleine Programmierer, die einmal irgendwas irgendwie machen, und glauben eine HSDB brauchen sie jetzt unbedingt, weil es gerade Hype ist, und für die ist es natürlich spannend kostenlose Lösungen zu machen. Mit ein bisschen Nachdenken kommen die da drauf, dass sie häufig auch die gleiche Leistungsfähigkeit mit einem mem-cache open drauf machen. Also es gibt einen key-value-store, der auch HSDB, soweit ich HSDB [inaudible], oder verwaltet jetzt Objekte im Hauptspeicher.

Und wie gesagt, das was die SAP damals gemacht hat war eine sehr genaue Analyse der Anfragemuster für die Applikationen, die ihnen am meisten Geld gebracht haben. Und daraufhin haben sie jetzt festgestellt, dass das meiste Geld beim Kunden nicht in die SAP geht für eine Datenbank, sondern zu Oracle oder zu IBM. Und das hat SAP gestört. Die hat gestört, dass bei 5€ Investment nur 2€ bei ihnen landeten, und 3€ an IBM oder Oracle für eine Datenbank gegangen sind. Und da haben die sich gefragt: Hör mal her, warum sollten wir das zulassen?

R [23:54-23:56]: Und warum war das so?

E8 [23:56-24:20]: Warum das so war? Aufgrund der Fokussierung der Unternehmen. Die SAP hat sich auf den Enterprise-Markt fokussiert und IBM und Oracle auf andere Märkte. Die SAP hat auch eine eigene Datenbank, aber diese eigene Datenbank zog halt nicht. Das war MaxDB. Aus der hat sich später dann MySQL entwickelt.

R [24:20-24:24]: Solche Informationen helfen mir auf jeden Fall auch sehr viel weiter.

E8 [24:24-24:27]: Die stehen aber auch in dem Text, beziehungsweise bei Golem drin.

R [24:27-24:29]: Welchem Text bitte? Ah, ja...

E8 [24:28-24:30]: Den Text den ich bei Golem hatte.

R [24:30-25:09]: Ja, den konnte leider bisher nur überfliegen. Ich habe das nicht geschafft. Dann schaue ich mir den nachher auch noch einmal ganz genau an.

Dann jetzt wird es ein bisschen allgemeiner. Nicht mehr nur auf RIMDB, sondern an sich diese Schlagwort Big Data Analyse. Und zwar wüsste ich gerne von Ihnen, welche gesetzlichen Regulierungen die Entwicklung von Big Data Analyse positiv oder negativ beeinflussen könnten.

E8 [25:09-25:13]: Hä? Was erwarten Sie denn jetzt als Antwort?

R [25:13-25:39]: Es könnte zum Beispiel sein, dass es gewisse Einschränkungen gibt durch Datenschutz, dass sich die Potenziale, die da drin stecken, nicht so entwickeln können wie man es gerne hätte, und wie sie vielleicht auch nützlich wären. Andererseits könnten auch gewisse Bereiche gefördert werden.

E8 [25:39-25:46]: Ah, das ist diese Datenschutz-Frage.

R [25:46-26:01]: Oder vielleicht auch im Bereich auf Standardisierung von Datenquellen... Das ist eine ganz offene Frage.

E8 [25:59-29:52]: Da kann ich nicht... Es gibt Immanuel Kant. Und der hat einmal ein bisschen was über die Ethik erklärt. Und die Ethik in Deutschland basiert unter anderem auf der Ethik, die Kant gegeben hat. Insofern basiert auch unsere Gesetzesgebung auf dem was Kant und das Römische Recht zum größten Teil gesagt haben.

Und die Schwierigkeit, die wir derzeit haben, das ist eine ganz andere. Die Schwierigkeit heißt, die Jurisdiktion und die Exekutive Probleme haben, die immer wieder neuen Begriffe in der IT genügend gut zu abstrahieren. Das heißt also, man muss eigentlich sehr sehr gut Richtern, also auch Leuten in den Ministerien, und das ist im Wirtschaftsministerium sehr sehr positiv, auch im Forschungsministerium, die haben da einen relativ guten Durchblick, begreiflich machen, welche Chancen und Risiken Technologien geben. Und dann wird sich das auf die entsprechende Gesetzgebung auch auswirken. Und dann gibt's natürlich Gesetzesentwürfe, die von der deutschen Lobby geprägt werden, und hier ist abzuwägen: Gewinnen wir am Standort Deutschland, wenn wir bestimmte Regularien nach unserem Wertesystem freier gestalten? Oder verlieren wir? Und das ist eine politische Abschätzung. Das ist eine Abschätzung inwieweit sozusagen Arbeitsplätze gefährdet sind.

Zum Beispiel wurde ja neulich die Akquise von [inaudible] zum Glück durch deutsche Konzerne erfolgt. Die erfolgte nicht durch das amerikanische Unternehmen Uber. Insofern können wir es also als Deutsche für unseren Automobilmarkt, der einen der vier Grundsäulen der deutschen Wirtschaft ist, Regularien erlassen, die also die Kartenbenutzung entsprechend steuern.

Und ansonsten ist der Aspekt Recht und Daten sehr vielfältig. Da gibt es eine sehr große Studie dafür, da steht auch mein Name drauf, das kann ich Ihnen mal schicken. Das wichtigste ist hier die Unsicherheit im Bereich Urheberrecht. Das geht zurück auf den Fall, dass also jemand einem eine Leinwand und Farben verkauft, dann

malte jemand, der es gekauft hatte darauf ein Bild, und dann kam derjenige, der die Leinwand verkauft hatte und sagte, das Bild gehört mir. Und teilweise ist es immer noch so, dass die Rechtsprechung dementsprechend ist. Das heißt also: Wie viel kreative Schöpfungshöhe müssen Sie in die Veränderung eines Datums herein stecken, dass das Datum Ihnen also nicht mehr urheberrechtlich zugerechnet werden kann, weil Sie es verkauft haben.

Das zweite ist Haftungsrecht: Sie verkaufen mir drei Kilo Gesundheitsdaten, und bei mir sterben dann zwei Kinder: Kann ich Sie verhaften, oder nicht? Dann haben Sie als nächstes das Thema Insolvenzrecht: Meine drei Kinder sind gestorben, Sie sind aber Pleite gegangen: An wen richte ich mich? Hat ihr Nachfolger dafür die entsprechende Konsequenz zu tragen? Das vierte, Punkt vier ist Roboterrecht: Der Roboter, den Sie mir verkaufen, fährt mir über den Fuß. Wer hat Schuld? Sie, der Verkäufer; Sie, der den Roboter design hat; Ich, weil ich meinen Fuß da hatte; der Roboter selber?

Und dann kommt erst Datenschutz. Erst dann. Weil Datenschutz leitet sich aus was ganz schwierigem aus dem Grundgesetz ab. Nämlich dem Recht auf Informationsfreiheit. Dem Recht auf Informationelle Selbstbestimmung. Und da haben einige einmal draus gemacht Datenschutz. Und das ist oft ziemlich was gegen [inaudible]. Und in der Problematik, in der wir rein rennen, ist es nur Punkt fünf. Auch wenn es einige andere gern hätten.

R [29:52-20:06]: Das wäre super, wenn Sie mir das schicken könnten, das würde mir sehr weiter helfen. Genau da ist es für mich sehr schwierig aus praktisch... aus den vielen... verschiedenen Quellen genau das Essentielle heraus zu ziehen.

E8 [20:05-30:12]: Es gibt nichts Essentielles. Es gibt ein großes Kuddelmuddel mit verschiedenen Lobbyorganisationen.

R [30:12-31:07]: Ja. Dann geht's im nächsten Punkt um die Verbindungsnetzwerke. Hier das Schlagwort Internet of Things [IoT], das ja sehr gerne und viel benutzt wird. Können Sie sich da vorstellen, dass... Also prinzipiell erstmal als Grundannahme von mir aus: Das IoT, diese ganzen Sensoren, sind natürlich eine große Datenquelle. Und könnte es passieren, dass das Netzwerk sich zu langsam entwickelt? Könnte sich das als Problem entwickeln? Oder sagen Sie: Das ist irrelevant, weil das ist sowieso, innerhalb von Unternehmen sind die meisten Daten, Netzwerk hat damit nicht viel zu tun.

E8 [31:07-33:25]: Also Netzwerk... Es gibt ja Studien darüber, was am schnellsten wächst. Und das Netzwerk, die Geschwindigkeit, der Bytes, die wir pro Sekunde über den Leiter bekommen, wächst leider am langsamsten im Vergleich zum [inaudible]-Speicher. Das Netzwerk ist natürlich besonders gravierend bei der ganzen Geschichte, und das Bottleneck.

Es gibt bestimmte Gebiete in Deutschland, in Flächenländern, die wirklich in der Tat noch sehr schlecht angebunden sind im Vergleich zu anderen Gebieten, in Hauptstädten. Und das ist natürlich dann für die Industrie, die dort zuhause ist und die dann irgendwelche Glühbirnen oder Heizungsgeräte ausliefert, die dann transaktional sehr viele Daten dann dahin schicken, doof. Ich glaube aber trotzdem nicht, dass wir uns in Afrika befinden, sondern in Deutschland, das heißt also: Es muss da Anstrengung natürlich geben, denn Breitbandausbau zu beschleunigen, aber es gibt viele andere Bausteine auch, die wir da haben in Deutschland. Und jetzt ist das natürlich wieder eine Frage des Preises. Natürlich hat jedes gute Unternehmen, wenn es es sich leisten kann, die Rechenzentren nahe bei Frankfurt, wo also der deutsche Internetknoten dran ist. Das hat natürlich seinen Preis. Und da sind natürlich einige Unternehmen, die das lokaler aufbauen, benachteiligt, weil sie transaktionale Volumen vielleicht nicht in Frankfurt haben dürfen in einem Rechenzentrum, sondern halt irgendwo anderes. Und dann kommt's zu Problemen, ganz klar.

Und das zweite Problem was wir hatten noch vor ein, zwei Jahren, war IPV6. Ich bin mir nicht ganz sicher, ich glaube, dass mittlerweile die meisten Unternehmen Router und Endgeräte bereitgestellt haben, die IPV6-fähig sind, und damit hat man jetzt also mehrere Größenordnungen mehr Adress-Spielräume. Ansonsten wäre das ein Problem geworden. Ansonsten hätten Sie nur 256 noch vier IP-Adressen gehabt, und damit kann man natürlich nicht so viel machen.

R [33:26-33:39]: Und sehen Sie da irgendwelche Trends? Zum Beispiel: Kommen da auch Entwicklungen eher aus der Wirtschaft um aufzuholen, und vielleicht Leitungen selber zu legen? Ist Ihnen da etwas bekannt?

E8 [33:39-35:46]: Was man jetzt schon sieht ist: Das größte Problem, das wir ja haben, ist dass sie ja kein Backbone... infrastrukturell keinen Backbone Hersteller in Europa haben. Also wir kaufen entweder bei Huawei

und lassen vom chinesischen Geheimdienst abhören, oder wir kaufen bei Cisco und dann ist es irgendjemand anders. Und Siemens hatte ja mal eine Router-Sparte, [inaudible] war das glaub ich [inaudible], das hat aber nicht so gut funktioniert. Und insofern finde ich es schon gut, dass jetzt also Nokia, der ehemalige Reifenhersteller und Handyhersteller auch wieder diese Infrastruktur bereitstellen will für's Backbone. Wie weit die jetzt sind, und wie weit sie sich Markt-Shares da erobern können weiß ich nicht. Aber es ist auf jeden Fall für Europa erstrebenswert, dass man bei einem vertrauten Partner, in diesem Fall Finnland, einfach Chips einkaufen kann, und man kann ihnen vertrauen, und hat nicht die Befürchtung, dass die Daten dann bei irgendwelchen Geheimdiensten auf der Welt landen. Ich glaube, das ist das wichtigste [inaudible], was momentan das Problem ist.

Weil ich mein, wenn wir transaktionale Daten erzeugen, dann ist das natürlich wiederum ein Abbild der Welt, und wir haben uns vorhin darüber unterhalten wie im Leben bezahlt wird. Und wenn Sie jetzt Modelle haben, die das Abbild der Welt charakterisieren, können Sie natürlich sehr schön Vorhersage-Modelle bringen: Jeder, der also diese Abbilder der Welt hat, hat eine Kopie der Welt, und kann Vorhersage-Modelle machen, die exklusiv sind, kann diese vertrieben. Und leider Gottes gehen halt die meisten Datenströme dieser Welt derzeit entweder nach Asien oder nach Nordamerika. Das ist ein strategischer Nachteil für Europa.

R [35:46-36:18]: Dann sind wir schon ganz fix beim letzten Punkt: Da geht es auch wieder um IoT, und zwar diesmal im Bereich Sensoren. Es gibt ja diese ganzen Vorhersagen, zum Beispiel von Gartner, dass das explosionsartig sich vermehren wird, und das IoT ist gerade ganz oben auf dem Hype Cycle. Wie sehen Sie das? Glauben Sie, das ist ein Hype? Oder das wird wirklich so passieren? Beobachten Sie da selber Trends?

E8 [36:19-39:20]: Nochmal, das ist die gesamte Frage, die gesamte Frage heißt halt: Wie werden Menschen in Zukunft Geld verdienen wollen, und es geht darum wer kann exklusive Daten, oder exklusive Modelle erzeugen. Also erstmal braucht man exklusive Daten, um aus den maschinellen Lernverfahren exklusive Modelle erzeugen zu können, die halt dann vertrieben werden können. Was man exklusiv hat, ne, wenn die Informationen so hochgradig stimmen würden, kaufen Sie morgen IBM Stocks würden Sie also viel Geld dafür bezahlen. Oder die Informationen halt wie man Autos baut, die nicht gegen den Baum fahren, das sind Exklusivmodelle. Oder das exklusive Modell bei Google, was Ihnen die Werbung schaltet, und so weiter, das geben die ja nicht raus. Oder bei Siri, das exklusive Modell was Ihnen sozusagen die Sprache erkennt. Und das machen die halt ganz gut. Und um diese exklusiven Modelle zu bauen benötigt man Rohdaten, möglichst frisch in möglichst großer Art und Weise. Und diese Rohdaten werden dann mit geheimen Algorithmen – aber so geheim sind die nicht, der Algorithmus ist meist eher das unproblematischste – zu Modellen verarbeitet und aktualisiert. Die müssen also möglichst frisch sein.

Und jetzt ist die Frage: Wenn also die Rohdaten der Rohstoff sind um das Produkt, das Modell, was sie als Dienstleistung anbieten, zu verarbeiten, dann ist die Frage: Wie produzieren Sie diese Rohdaten? Und derzeit haben Sie Rohdaten durch ihr Handy produziert, das weiß also zum Beispiel, also Google weiß sehr genau ob sie Fahrrad fahren oder ob sie nicht Fahrrad fahren, weil Google weiß wie schnell Sie sich bewegen, und wie schnell das Ding geschüttelt wird. Und bei 20km/h, oder 18, und viel Schüttelei kann man drauf kommen, dass Sie gerade Fahrrad fahren. Besonders wenn Sie das [inaudible] machen. Da fahren Sie nicht so lange 20km/h. Und das Klickverhalten im Browser wird halt für Sie ausgenutzt, dass sie besseres Werbeverhalten bekommen, oder die Suchanfragen, oder wie schnell Sie die Maus bewegen. Also da gibt es ganz viele Sensoren schon, die wir eh schon benutzen. Der größte Sensor ist Ihr Handy.

Und jetzt möchte man das Auge aber, dieses Gehirn, also Modelle sind das Gehirn, und das Gehirn braucht ja mehr Augen und Tasten und so weiter, möchte man das Auge halt größer gestalten. Warum sollte man das nicht tun, wenn die Preise für Sensoren sinken? Warum sollte man das nicht tun, wenn man das einer Bevölkerung verkaufen kann als das bringt dir was?

Jetzt kommen wir zur deutschen Bevölkerung, die mit Kant aufgewachsen ist, und stellt plötzlich fest: Die sind da kritischer. Und das ist das Problem. Also es kommt wiederum auf die Ethik an, es gibt ja wie gesagt im amerikanischen und asiatischen Raum viele Menschen, denen es nicht sehr gut geht, die gerne im Mittelstand aufwachsen wollen, und die total bereit sind sich mit Sensoren zu bepflastern und damit durch die Gegend zu laufen wie ein Weihnachtsbaum.

R [39:21-39:30]: Aber glauben Sie, dass das auch wirklich weiterhin so ein großes Problem bleibt in Deutschland? Oder wir uns da auch angleichen?

E8 [39:30-39:31]: Haben Sie Kinder?

R [39:31-39:32]: Nein.

E8 [39:32-39:44]: Aber Sie sind wahrscheinlich nach Google geboren, Millennium-Generation sind Sie. 1995 oder irgendwie sowas.

R [39:44-39:50]: Nein, neunzehnjähriger. Ich bin noch gerade ein „Digital Native“.

E [39:50-40:38]: OK. OK. OK. Aber die nächste Generation ist schon Millennium, und die übernächste... Mein Großvater ist noch aufgewachsen und fand die Errungenschaft der Glühbirne ganz toll, und meinte man sollte nicht Radio hören, sondern Lesen. Das heißt also: Jede Generation bekommt neue Werte mit. Und insofern ist das die große Frage für unsere... Ist wiederum eine philosophische Frage, ist die Frage für unseren Raum in Deutschland: In wie weit können wir es uns noch leisten, eine bestimmte Ethik aufrecht zu erhalten. Und ich würde mir wünschen noch sehr, sehr lange.

R [40:39-40:50]: Gut, ich denke das ist ein wundervoller Abschluss. Dann sind wir nämlich auch schon durch alle Themenblöcke durch.

[40:50-50:11]: [closing conversational remarks and explanation of further research process]

Attachment C: Interview Transcripts and E-Mails Phase R3

Fact Sheet for Phase R3

What is a Key Driver?

Key drivers are those influencing factors that were found to be most relevant as influencers to the future development of the objective (i.e. in this case of a technology)

What is a Projection?

A projection is the end stadium of a possible future development for one key driver. Normally, several projections are imaginable for each key driver. The projections for one key driver can differ qualitatively, and are not necessarily just a good, neutral and bad scenario of the future development.

Example:

Key Driver A: Downpour

Projection A1: "It snows."

Projection A2: "It rains."

Projection A3: "It hails."

What is the Consistency Matrix?

In the consistency matrix, two projections for different key drivers are always looked at simultaneously, and rated on a scale from -3 (very inconsistent) to +3 (very consistent).

The degree of consistency describes how possible it is for both projections to take place at the same time.

If the two projections are very consistent with each other, they are a good match, if the projections are very inconsistent they exclude each other from happening simultaneously.

Example:

Key Driver A: Downpour → Projection A1: "It snows."

Key Driver B: Temperature → Projection B1: "It's 30°C."

Projection A1 and projection B1 are highly inconsistent (-3).

Key Driver A: Downpour → Projection A1: "It snows."

Key Driver B: Temperature → Projection B2: "It's -5°C."

Projection A1 and projection B2 are highly consistent (+3).

Transcript E1

R [00:00-00:50]: Das war praktisch die Vorbereitungsstufe, daraus, und aus anderen Interviews haben wir dann sogenannte Driver, also Einflussfaktoren kristallisiert, und die wurden dann in ein Netz eingegeben und mit verschiedenen Einflüssen verbunden, und daraus haben sich dann sechs Key Driver ergeben. Die Key Driver stehen immer hier vorne, da ist einmal demand of real-time response and service acceleration. Dann die Kosten, dann open-source, also relationale open-source Lösungen, Regierungseinflüsse, das Netzwerk und die Sensoren.

E1 [00:51-00:53]: OK, was ist mit Netzwerk konkret gemeint, also jetzt das technische Netzwerk, oder...

R [00:53-02:24]: Genau, also die Entwicklung. Da geht's mehr in Richtung Internet of Things [IoT], dass z.B. auch die ganzen Sensoren ans Netzwerk angeschlossen sind, z.B. Entwicklung von openWifi oder 5G. So in die Richtung.

Und dann habe ich da wiederum Interviews geführt, und wir haben aus diesen Interviews für jeden der sechs Key Driver Projektionen entwickelt. Eine Projektion ist ein zukünftiger Zustand mit dem Zeithorizont 2025. Und sind eben mögliche Zustände – muss dann nicht immer so sein: es steigt, es sinkt, es bleibt gleich, sondern möglichst sollten die auch qualitativ unterschiedlich sein. D.h. manchmal klingen die dann erst ein bisschen gleich, aber haben einen anderen Hintergrund. Und worum es heute geht ist, dass wir alle diese Projektionen... mussten wir gegeneinander abgleichen auf ihre Konsistenz, weil daraus sich dann am Schluss die Zukunftsszenarien ergeben. D.h. wir haben... Und die werden dann bewertet auf -3, sehr inkonsistent, auf +3, sehr konsistent. D.h. -3 wäre die passen überhaupt nicht zusammen, das kann gar nicht sein, dass die gleichzeitig in einem Szenario auftreten. Und +3 wäre das bedingt sich perfekt, es passt perfekt zusammen.

E1 [02:24-02:30]: Also ihr habt jetzt hier auf der Horizontalen die Key Driver eingetragen, und dann auf der Vertikalen auch nochmal?

R [02:31-03:01]: Genau. Und dann werden die miteinander abgeglichen. Vor allem der Themenbereich open-source hat uns ein bisschen Probleme bereitet, weil wir da nicht wirklich viele Verbindungen gesehen haben, und deswegen steht hier überall 0. Vor allem das würden wir dann mit dir nochmal absprechen. Vielleicht ist es am besten wenn du dir die eben einfach mal durchliest. Und falls du irgendwo fragen hast oder etwas nicht verstehst kann ich das dir dann noch erklären. Und rufe das solange mal hier auf.

[03:01-03:51]: Pause

R [03:51-04:13]: Und praktisch die Grundlage ist ja, für was wir die Szenarien bauen, die sollen zeigen wie die Entwicklung von relationalen in-memory Datenbanken [RIMDB] aussieht. Und über diese Voranalyse haben wir dann diese Driver festgestellt.

[04:14-04:19]: Pause (Device off for several minutes)

E1 [04:19-06:20]: Das geht so in die Richtung des Data-Warehousings, nicht? Ich sag mal so: Früher, oder aktuell eigentlich immer noch, sieht ja eher eine Data-Warehouse-Architektur so aus: Wir haben verschiedene Quellsysteme, dort werden Transaktionen verarbeitet wie jetzt im CRM oder im normalen Enterprise Resource Planning, wie bei SAP. Und man schaufelt halt diese Daten in gewissen Zyklen in so ein Data-Warehouse rein um da jetzt Analysen zu machen. Weil die meisten Datenbanken [DB] sind entweder transaktionsorientiert oder eher für Analysen. [inaudible]

Und weil man eben versucht mit in-memory DB das Problem zu lösen, dass du mit in-memory DB, also gerade die SAP mit HANA den Anspruch hat, dass du sowohl den OLTP, also den transaktionsorientierten Bereich, als auch OLAP, den Analysebereich in einer DB machen kannst, also dass du da nicht mehr ein extra Data-Warehouse für brauchst. Und vor allem, dass du keine Zyklen mehr dazwischen hast, weil du hast ja immer einen Transaktionsaufwand zwischen der DB und dem Data-Warehouse, und den Quellsystemen. Du hast eigentlich, ich sag mal, Analysen direkt auch in den Quellsystemen zu machen, dass... den Anspruch hat eben auch eine SAP damit, das eben auch mit in-memory Technologien zu lösen, vor allem mit HANA. Deswegen stellen die auch ihre ganze Suite, also das ERP-System, CRM-System, das früher auf klassischen DB lief, stellen die ja komplett um auf HANA, dass das eben auch auf HANA läuft. Und dass du eben auch OLAP, also die Analyse, das Analyse-Processing auch in der DB machen kannst, ohne jetzt eben ein Data-Warehouse...

Also ein Data-Warehouse brauchst du schon irgendwie, sagen wir mal, in der Architektur, als Schicht um da jetzt Analysen zu machen, oder du hast einen Cube oder irgendwie ein Drill-Down machst in die Quartalszahlen auf ein Jahr, dann geht's weiter rein, pro Quartal pro Woche, und diese klassischen Analysen, dazu brauchst du immer noch diese Tools, aber du hast praktisch als DB eine Technologie darunter, und musst das jetzt nicht von zwei getrennten DB speichern.

R [06:20-06:24]: OK. Und dann ist das... passt dann nicht so dazu?

E1 [06:24-06:50]: Also ich hab's jetzt... Also das ist, genau, grad in-memory ist in diesem Bereich OLAP und OLTP, also gerade für Analysen, ein wichtiger Punkt. Das hatte ich jetzt hier in den Key Driver nicht so gesehen. Oder wo das vielleicht drin steckt. Also dass man eben versucht OLTP und OLAP, also diese beiden Welten miteinander zu verbinden. Das passt vielleicht hier oben ganz gut, also demand of real-time response and service acceleration. Ja.

R [06:50-07:17]: Also die sind natürlich schon auch bisschen high level diese Driver. Wir haben auch... Ich denke das sollte schon da drunter gefasst werden. Und dann war halt auch die Frage: Was ist eigentlich real-time? Und wie real-time kann's noch werden, kann sich das noch immer weiter steigern? Und deswegen haben wir da auch so ein bisschen....

E1 [07:17-08:12]: Ja klar. Unsere Definition, oder auch die Informatik, sagt... Also in der Cloud gibt es eigentlich kein real-time. Es gibt wenn dann nur near-real-time. Weil real-time machst du entweder nur wenn du im Flugzeug bist, ein geschlossenes System, ein geschlossenes Netzwerk, oder auf einem Device, da hast du real-time. Aber in der Cloud, wenn du da die DB rein bringst, hast du immer eine Latenz dazwischen. Das ist eine Definitionssache. Also für uns ist das, was sich in einer Cloud abspielt, immer near-real-time, weil da ist wenn auch nur im Sekundenbereich immer eine Diskrepanz dazwischen.

Ich weiß jetzt nicht, das ist ja für die Abschlussarbeit, in wie weit für die Bewertung... Also ich denke mal, OLTP und OLAP, wenn die als Begriffe irgendwo fallen kann ich mir vorstellen, dass das wichtig ist, dass man das vielleicht nochmal mit rein bringt.

Zum anderen weiß ich jetzt nicht wie weit wichtig real-time, near-real-time, inwieweit man da jetzt auf die Differenzierung achtet. Das ist immer eine Definition...

R [08:13-09:11]: Ich glaube wenn ich da... Das soll ich dann hier schon noch mal definieren, dass es da verschiedene Möglichkeiten gibt, das da auszudrücken.

OK. Dann können wir gleich mal in die Konsistenzmatrix rein gucken. Ich würde sagen wir fangen erstmal mit dem open-source Block an, und dann können wir schauen wie viel Zeit wir noch haben.

Die ersten zwei Projektionen, die wir angucken, sind open-source und kommerzielle nebeneinander, und eine Reduktion der hardware-Kosten. Und da hatten wir überlegt: Wie könnte das zusammen hängen? Könnte es da einen Zusammenhang geben, dass die hardware-Kosten sich reduzieren und dass sowohl open-source als auch kommerzielle Lösungen gleichwertig auf dem Markt sind?

E1 [09:12-10:20]: Also ich würde da jetzt... Also erstmal generell denke ich, die hardware-Kosten werden sich immer reduzieren, also das ist ein ungeschriebenes Gesetz. Oder wie heißt das Law... Morsh Law... Also ich sag mal, das wird sich weiter verbreiten, weil früher war in-memory, und der RAM war teuer. Ein Speichermedium wird immer günstiger sein, das jetzt einen direkten impact auf open-source oder commercial hat weiß ich nicht. Das gilt insgesamt für beides, dass die Verbreitung steigen wird, also dass das der Einfluss ist. Aber wenn man mal auf dem Markt schaut: Es hat sich der open-source in-memory Bereich nicht wirklich weiter entwickelt. Dort ist eine SAP mit HANA ganz stark, dort ist eine Oracle mit ihrer TimesTen relativ stark, dort ist eine Microsoft mit dem Hackathon und dem SQLServer relativ stark. Also ich hab bisher noch kein Projekt gesehen wo wirklich open-source in-memory DB, wie man es jetzt von HANA kennt, wirklich angewandt werden.

R [10:21-10:39]: Das heißt: Hier gleich der nächste Punkt, die open-source übernehmen den Markt – natürlich... Darum geht es, dass es qualitativ ist, vielleicht ist es nicht so wahrscheinlich, aber theoretisch könnte es sein.

E1 [10:39-11:01]: Also ehrlich gesagt sehe ich hier... Also wenn es jetzt um eine Expertenmeinung geht sehe ich das nicht, dass open-source den Markt im in-memory Bereich übernehmen wird, weil dort die Player einfach zu

stark sind als Begründung, weil die SAP, Microsoft ihre Produkte rein pushen und vermarkten, das auch direkt mit den... ihren bestehenden... Bestandsprodukten verbinden.

R [11:02-11:08]: Das heißt da wäre auch keine Verbindung zu den Kosten von hardware in irgendeiner Form.

E1 [11:08-11:14]: Zu in-memory insgesamt jetzt, aber nicht in der Unterscheidung open-source oder...

R [11:14-11:56]: Das ist eben die Schwierigkeit... Das heißt, das können wir eigentlich... Im Grunde: Vielleicht ist es auch einfacher, wenn wir nicht alle einzeln durchgehen sondern überhaupt einfach mal gucken: Könnte zwischen irgendeinem dieser Faktoren eine Verbindung zu open-source versus kommerziell passieren? Also die hardware-Kosten, die Projektkosten sinken durch mehr Erfahrung... Cloud wird populärer, da haben wir hier einen ganz leichten Zusammenhang, weil wir dachten, naja, das ist jetzt nicht so direkt verbunden, aber es könnte es leichter machen.

E1 [11:56-12:39]: Also Cloud, genau... Also Cloud verbreitet sich mehr, und das wird aber eher nicht den open-source DB zu Gute kommen, sondern den kommerziellen, einfach aus dem Grund: Wer ist der Cloud-Player? Also SAP hat eine HANA-Cloud, eine Microsoft – kennt ihr Azure, schon mal davon gehört? Also das ist die Cloud Plattform von Microsoft. Azure hat von Microsoft. Azure hat seit [inaudible], und das heißt sie bringen natürlich auf ihren Azure-Plattformen ihre Produkte dann drauf. Oracle genauso, Oracle hat auch eine Cloud. Und ja... Und die bringen da natürlich auf ihre Cloud-Plattformen die DB, die sie auch selber entwickelt haben. In-memory DB.

R [12:40-12:49]: Das heißt, das spielt sich praktisch gegen open-source, die Cloud, und fördert, dass es irrelevant bleibt.

E1 [12:49-12:51]: Genau.

R [12:51-12:55]: Würdest du sagen ein sehr starker Zusammenhang?

E1 [12:55-12:56]: Zwischen was jetzt, zwischen open-source...

R [12:56-13:03]: Genau, und dass Cloud immer üblicher wird.

E1 [13:04-13:14]: Ja, also ich... Wie gesagt, aus meiner Sicht wird sich die nicht-open-source DB, also die kommerziellen DB weiter verbreiten, und stärker in den Markt kommen als...

R [13:14-13:54]: OK. Dann auch gleich noch ein interessanter Punkt sind diese Hybrid-Systeme, damit meinen wir, dass das System weiß welche Daten auf welche Art von Speicher, oder auf welcher Art von System gespeichert werden. D.h. das war dieses hot, warm und cold data approach. D.h. dass du manche vielleicht gar nicht auf so schnellen Systemen brauchst, vielleicht nur auf irgendwelchen schnelleren disks... Und ob es da einen Zusammenhang gibt? Zum Beispiel, dass das open-source fördert? Oder...

E1 [13:54-14:00]: Also das sehe ich neutral. Neutral, genau.

R [14:01-14:11]: Das ist im Grunde ähnlich wie der erste Punkt hardware-Kosten. Die hardware wird zwar günstiger, aber auch effizienter, das heißt es stagniert im Grunde.

E1 [14:12-14:15]: Also auf open-source versus commercial ist es auch neutral.

R [14:15-14:38]: Hat nichts damit zu tun, OK, dann haben wir's ja gut eingeschätzt. Dann vielleicht als nächsten Block die Regierung, Regularien.

E1 [14:39-16:37]: Da kenne ich mich ehrlich gesagt nicht so aus. Also, ja... Machen wir erstmal. Ich sehe da, wenn du in diese Matrix schaut, zwischen Regulatoren und wenn du hier Cloud, da gibt's natürlich einen Zusammenhang. Und zwar wird...

Also in dem Sinne, dass der Cloud-Markt in Deutschland durch dieses Safe Harbor Abkommen, was ihr wahrscheinlich gehört habt, das ja abgekündigt wird, gibt's da einen impact, sodass der Cloud-Markt in Deutschland vermutlich weiter wachsen wird. Und dass eben eine Microsoft mit ihrer Cloud-Plattform dahin geht... Es gibt auch einen Medien-Artikel wenn ihr darauf [inaudible]. Also Microsoft wird das Thema Azure als German-Cloud, also ihre Cloud-Plattform mit der in-memory DB, auch in Deutschland boosten. Also auf der CBit

auf einer T-System Infrastruktur... [inaudible]. Gibt auch einen Galileo-Artikel dazu. Das ist jetzt nicht wissenschaftlich, was die da machen, aber kann man... Und die T-Systems, oder auch die Deutsche Telekom wird dann als Data Trustee auftreten, also die Microsoft selber hat auf ihre Plattform dann gar keinen Zugriff mehr, sondern nur durch die Telekom als Data Trustee.

Und das hat natürlich schon die Regulatoren im Zusammenhang mit Safe Harbor Abkommen, da findet man auch z.B. Sachen im Netz, dass der Datenaustausch jetzt in die Vereinigten Staaten dadurch eingeschränkt wird. Und die Daten praktisch in Deutschland bleiben, oder bleiben müssen, wird sich halt der Cloud-Markt da denke ich in Deutschland weiter stärken in der Richtung. Was man eben auch daran sieht, dass eine Microsoft da hin geht und sagt wir hosten das jetzt komplett in Deutschland. Die werden zwar auch in den anderen Lokationen ihr Rechenzentrum klar bestehen lassen, aber die bieten eben das Thema Cloud und in-memory DB eben wirklich für den deutschen Markt isoliert in einem Rechenzentrum an.

R [16:37-16:57]: OK, das heißt, würdest du... Wir haben hier nämlich eine negative Verbindung zwischen, dass es abgeschottete Systeme gibt, wo Deutschland auch ein bisschen ein Spezialfall ist, das heißt du würdest eher sagen es ist positiv bedingt, dass Cloud dadurch wächst, weil das System abgeschotteter ist? [inaudible]

E1 [16:57-17:06]: Also man muss differenzieren: Cloud-Markt in Deutschland, oder Cloud-Markt in Europa... Also man muss da regional differenzieren.

R [17:06-17:26]: Ja, das ist insgesamt so ein bisschen schwierig. Wir haben hier einen Split der Regularien, also dass es z.B. in Deutschland alles geschlossener ist, und in den USA und Asien alles viel offener. D.h. da würdest du auch einen positiven Zusammenhang sehen für Cloud...?

E1 [17:26-18:09]: Also in Deutschland. Das muss man wahrscheinlich dazu sagen. Also insgesamt, wenn man es global betrachtet, natürlich negativ. Weil: Je mehr Einschränkungen es gibt, desto weniger Freiheiten, desto weniger wächst das. Man muss da vielleicht eher... Das ist hier schlecht abbildbar, aber vielleicht in Text-Form darauf eingehen welchen Impact das gegebenenfalls auf Deutschland auch hat. Also der Cloud-Markt wird weiter wachsen, das ist ganz klar, nur eben mit gewissen Einschränkungen, und eben auch mit impact... dass in Deutschland beispielsweise dann Sachen gehostet werden.

R [18:10-18:30]: Und wenn wir nochmal auf open-source gucken. Auch solche Regularien wie dieses Safe Harbor Abkommen, wie das darauf einwirken könnte – hat das damit irgendetwas zu tun überhaupt?

E1 [18:31-19:13]: Nein. Ich kann mir nur höchstens vorstellen, weiß ich nicht, also dafür kenne ich... Also ob das open-source ist oder nicht... Ich kann mir höchstens vorstellen, weiß ich nicht, irgendwelche Regularien... Ich mein, was da jetzt mit der NSA passiert ist, da müssen wir nicht drüber reden, das kennt ja jeder, dass es da Regularien gibt, das vielleicht verpflichtet, Anbieter wie eine Microsoft, weiß ich nicht, oder das ist jetzt nur theoretisch, rein exemplarisch, andere Anbieter verpflichtet irgendwelche Sachen da einzubauen, damit Behörden da Zugriff drauf bekommen. Und da ist aber ein bisschen... Früher hätte man das als Verschwörungstheorien abgetan, aber heute ist es ja leider Realität...

R [19:12-19:14]: Da weiß man leider es stimmt...

E1 [19:15-19:54]: Und vielleicht eher da wird's vielleicht dem open-source Markt gut tun, weil open-source ist open-source, da kann jeder in den Quelltext rein schauen, wenn es um sowas geht in Richtung was durch Snowden und so weiter ans Licht gekommen ist, Daten gespeichert werden, oder gerade es in der Cloud Zugriffsmöglichkeiten gibt. Wenn man eben open-source anbietet, dann kann man selber rein schauen und gucken ob es da irgendwelche Lücken gibt, oder nicht.

Aber mit Safe Harbor, das hat da nix mit, das wäre eher praktisch ein anderer Punkt.

R [19:54-20:45]: Aber hier haben wir nämlich auch, dass die... als eine Projektion, dass zwar alles offen ist bezüglich Sharing of Data und so weiter, aber die Regierungen den Profit daraus ziehen wollen aus den Daten. Das heißt das wäre dann schon ein stärkerer Zusammenhang... positiv... Und hier dann vielleicht eher negativ. OK. Cool.

Und diese Projektion hier, dass open-source als Vehikel für andere Systeme, wie zum Beispiel ERP-Systeme benutzt wird? Würdest du da irgendeinen Zusammenhang sehen?

E1 [20:45-21:58]: Also sehe ich ehrlich gesagt... dass wenn man sich den Markt anschaut... wird sich da in dem Bereich open-source sehr schwer tun, weil: Wer bietet ERP-Systeme an, für Enterprises die SAP, mit ihrem ERM- und ECC-System, die werden das Thema HANA da natürlich pushen als in-memory DB, was gibt es noch für CRM- oder ERP-Systeme, die Microsoft mit Dynamics-CRM, die werden auch ihr Hackathon, ihren SQL-Server als in-memory DB da mit rein bringen. Also dadurch dass der ERP-Markt sehr stark durch kommerzielle Anbieter gedeckt ist denke ich, werden sich da auch die kommerziellen DB in dem Bereich eher durchsetzen. Es gibt natürlich auch open-source CRM-Systeme wie SugarCRM, keine Ahnung, vielleicht, aber ich sehe insgesamt betrachtet, dass was jetzt der Schwenk von ERP-Systemen, CRM-Systemen auf in-memory Systemen, dass es eher den kommerziellen Anbietern zu Gute kommt, und die sich da durchsetzen. Weil die, die im Markt vertreten sind, auch ihre eigenen Technologien dann letztendlich verwenden, unterstützen wollen.

R [21:59-22:21]: OK. Dann haben wir noch zwei Blöcke, wo wir auch gar keine Einflüsse mit open-source zusammen gefunden haben, und zwar einmal das Netzwerk... Könntest du dir irgendwo eine Verbindung zwischen open-source versus kommerziell und wie sich das Netzwerk entwickelt vorstellen?

E1 [22:21-23:45]: Hm, also bei der open-source gibt es direkt auch keinen impact. Also was ein spannendes Thema wird, du hattest es ja schon angesprochen, 5G. Das ist natürlich nochmal ein Punkt, wo man ja einen Teil der Logik, oder sagen wir mal, ein Mini-Rechenzentrum, in die Basis-Station rein bringt. Also mit diesen sogenannten Cloudlets.

Heute ist es ja so du hast ein Device und ein Rechenzentrum, und dazwischen ist Connectivity, und die senden Daten da zwischendrin. Das Device kriegt dann da von dem Rechenzentrum irgendwann mal eine Antwort. Durch den 5G Standort wird es so sein um die Latenzzeiten möglichst zu verkürzen, was zum Beispiel für autonomes Fahren wichtig ist, dass ein Teil der Logik, die im Rechenzentrum aktuell läuft, auch in die Basis-Station rein kommt. Das heißt in jeder Basis-Station, was halt nur Connectivity darstellt, also irgendwie ein Turm wo ein Provider Mobilfunk anbietet wird eben ein Mini-Rechenzentrum sein, für alle Devices, die sich in dem Zuge befinden.

Und da wird's natürlich spannend inwieweit sich das Thema in-memory DB eben auch sagen wir mal als Teil des Rechenzentrums in den Basis-Station integrieren lässt. Also das ist noch zu früh zu sagen ob das eher dem open-source Bereich oder den kommerziellen DB zu Gute kommt.

R [23:46-23:48]: Das heißt da können wir auch 0 einfach lassen.

E1 [23:47-23:48]: Mhm.

R [23:49-23:47]: Gut. Und dann haben wir noch die Anzahl der Sensoren. D.h. das IoT und alles was dazu gehört. Also wir haben einmal unterschieden in den Projektionen als wirklich Sensoren, die auch alle ans Netzwerk angeschlossen sind, und tatsächlich dieses IoT sind, und dann nochmal: Es gibt zwar mehr Sensoren, aber das heißt nicht automatisch, dass das IoT ist, und alle sind verbunden.

Siehst du irgendeine Verbindung zu open-source tendenziell?

E1 [23:47-23:43]: Klar, also in-memory DB ja, aber ob es open-source ist oder nicht.... Dazu gibt es auch einfach zu wenig open-source in-memory DB, die wirklich produktiv eingesetzt werden können. Ich weiß nicht, mit welchen hattet ihr euch denn da mal beschäftigt?

R [23:44-25:02]: Also... So richtig populäre open-source haben wir nicht gefunden. Wir haben das rein genommen, weil ein Interviewpartner, oder zwei eigentlich das ziemlich betont haben, vor allem weil sie meinten die Kosten sind viel zu hoch für viele Kunden.

E1 [25:02-25:04]: Ja, das stimmt.

R [25:04-25:40]: Und dass es eine parallele Entwicklung in vielen Bereichen gab, also dass erst früher die kommerziellen Anbieter groß waren und dann ein Trend zu open-source kam, die dann halt einen hohen Industriestandard gesetzt haben, und dadurch Stück für Stück open-source der State of the Art geworden ist sozusagen. Und manchmal gibt es dann halt parallel open-source Lösungen und kommerzielle, und manchmal entwickelt sich da dann so ein System, dass open-source der goldene Standard ist und die business models eher um den... außenrum...

E1 [25:40-26:37]: Ja ok, also wenn man es so betrachtet kann man natürlich sagen, dass... Also ich sage mal für den, wenn man den IoT Bereich einmal nimmt, was ja eher unstrukturierte Daten sind, und einmal den classic

transaktionalen Bereich, wie Finance-Daten, wie CRM, ERP, die ja eher strukturierte Daten sind, da kann man schon vielleicht eine Tendenz mal sagen... Also ich denk mal so im Bereich strukturierte Daten und Transaktion wird sich in-memory eher für die kommerziellen auszahlen, weil die eben wirklich auch die ERP-Systeme herstellen, und darauf setzen. IoT ist natürlich recht offen und sagen wir mal, da kann man schon... Da könnte ich mir schon vorstellen, da gibt es aber wie du schon sagst zu wenig am Markt, aber dass sich im Bereich IoT, was eher offen ist, auch mit offenen Standards agiert und auch agieren möchte, wo sich auch Konsortien verschiedener Unternehmen zusammen tun, dass sich dort das Thema open-source vielleicht sogar stärker durchsetzen wird.

R [26:37-26:41]: Das heißt da können wir hier so einen leichten positiven Einfluss eintragen.

E1 [26:41-26:55]: Ja. Also wenn man eben diese Differenzierung zwischen IoT und Transaktionssystem zieht, ja...

R [26:55-29:05]: Ja, OK. [brief pause] Die sind dann immer genau gegenteilig. [brief pause]. OK, cool. Haben wir da schon mal ein paar Fragen geklärt. Dann vielleicht noch als anderen Punkt: Wir haben hier zweimal aufgeteilt mit einem regionalen Split. Einmal in Bezug auf die government regulations, und einmal in Bezug auf das Sensorenwachstum, denn unsere Theorie war: Wenn Europa, wenn das alles restriktiver ist, dann kann es auch sein, dass sich diese Sensoren nicht so wahnsinnig ausbreiten, und auch von... ja... von der Einstellung der Menschen her, dass Leute eben nicht so bereit sind überall Sensoren zu installieren. Haben wir aber leider auch ziemlich viele Felder wo es gar nicht einen so starken Einfluss hat. Vielleicht könnten wir die auch nochmal eben durchgehen.

Das ist einmal hier bei State of the Art ist... Genau, real-time response ist State of the Art geworden, und es gibt eine mittel bis starke Nachfrage. Könntest du dir da einen Zusammenhang vorstellen zwischen es gibt diese engen und... diese geschlosseneren und offeneren Systeme? Also hier wo wir es aufgeteilt haben meinen wir eben, dass praktisch diese real-time response wird gefördert oder hängt stark zusammen mit offeneren Systemen, und leicht mit geschlosseneren Systemen. Da wäre der Zusammenhang nicht so groß. Aber hier bei dieser Aufteilung haben wir irgendwie nichts...

E1 [29:05-29:55]: Hm, naja, also real-time response... Ich mein diese Anforderung kommt ja eher aus dem Unternehmens... unternehmensmotiviert, und... ja da ist man ja eh... Klar muss man gewisse Richtlinien betrachten, ob die vielleicht gar nicht so sehr mit einer Regierung, sondern auch Betriebsräte auflegen, ob man eben Auswertungen über Mitarbeiter in real-time haben möchte, überprüfen möchte Produktivität und solche Punkte... hm... Das ist irgendwie schwer zu sagen. Ja... Ob es da jetzt große... [brief pause]

R [29:55-30:39]: Also die grundlegende Idee war halt, dass diese Regularien schon die... einfach auch die Daten... welche Daten, und wie viele Daten ausgewertet werden, eingeschränkt wird, und dadurch, also vor allem wenn es um Personen geht. Und dadurch... ja, deswegen haben wir hier diesen positiven Zusammenhang mit den offeneren Systemen, z.B. auch da wo die Regierungen das ausnutzen, weil die davon selbst profitieren wenn möglichst viel da ist und möglichst viel auch in real-time ausgewertet wird. Und es ist natürlich jetzt schwierig wenn man sagt es gibt manche, die machen das, und manche nicht.

E1 [30:39-30:41]: Mhm. Wo hattet ihr das denn mit den geschlossenen Systemen?

R [30:41-30:51]: Hier. Und da haben wir dann so einen leichten Zusammenhang, weil wir dachten das heißt ja: Innerhalb der Systeme kann es sich ja trotzdem noch entwickeln...

E1 [30:51-32:08]: Genau, das meinte ich... Wenn man im Unternehmen ist, ja. [brief pause] Jaaa... Also wie gesagt, der demand kommt ja eher, sagen wir mal unternehmensgetrieben, und nicht getrieben von irgendwelchen Regierungen, sondern wo das Unternehmen eine Steuerungslogik braucht, schnellere Auswertung, und... wüsste ich jetzt nicht in wieweit da eine Regierung... Also die Sache mit dem Safe Harbor, wenn man das jetzt übergreifend sieht, kann es vielleicht einen negativen impact haben, aber... Und vielleicht die Betriebsräte oder sonst irgendwelche internen Stakeholder dort gewisse Beschränkungen machen was real-time Auswertung angeht, und wenn es mehr Beschränkungen gibt, dann ist klar, dann schränkt sich das ein, und wird eher einen negativen Aspekt haben, dass irgendwie gewisse Sachen verzögert sein müssen aufgrund irgendwelcher Regularien, und vielleicht anonymisiert... Was eben dann einen negativ impact auf das Thema real-time hat. Also ich würde es jetzt nicht zu stark negativ sehen, vielleicht leicht, oder sogar neutral...

R [32:08-33:00]: Genau, das meinten wir eben, wir haben jetzt hier die 1 genommen, also schon leicht positiv, weil wir auch dachten, dass... fördert das vielleicht nicht gerade stark, aber trotzdem wird sich das innerhalb der Systeme und mit gewissen anderen Daten, oder anonymisierten Daten... [inaudible]

OK, das haben wir schon angeschaut. Einen Moment... [brief pause]

Dann nochmal der regionale Split in die andere Richtung, und zwar bezüglich Sensorenwachstum. Da haben wir in Bezug auf Kosten überhaupt keine Zusammenhänge.

E1 [33:00-33:01]: Kosten jetzt...

R [33:02-33:27]: Also RIMDB-Kosten auf der Kundenseite, und dann haben wir hier ja verschiedene cases. [inaudible]. Hier genau. Also Kostenreduktion durch hardware, durch günstigere Projekt-Implementation, Cloud-Systeme...

E1 [33:27-34:16]: Also in die eine Richtung ist es denk ich positiv, weil sich das eben, sagen wir mal, die hardware-Kosten minimieren. Kannst du natürlich eine Vielzahl von mehr Sensoren und Devices auch verarbeiten. Also in dem Sinne wird praktisch... in dem Sinn, dass sich das hier, der impact, sagen wir mal die hardware und Kosten sich reduzieren für die hardware, das kann ja zum einen die hardware im Rechenzentrum auch sein für die in-memory DB, und die hardware selber für so einen Device, für so einen Sensor, wird das natürlich zur Folge haben, dass das einen impact hat auf die... also einen positiven impact hat, dass die Anzahl der Devices wächst, weil einfach mehr verarbeitet werden kann, weil die hardware eben günstiger wird.

R [34:16-34:34]: OK. Hier haben wir auch einen ganz starken positiven Zusammenhang. Das heißt: Auch wenn es regional aufgeteilt ist, insgesamt dann schon doch ein leicht positiver Einfluss. Oder Zusammenhang, kein Einfluss.

E1 [34:33-34:34]: Mhm.

[34:35-34:50] [brief pause]

E1 [34:50-34:55]: Ich glaub hier die relational... Warte mal, das ist, muss man das lesen von hier nach da, oder von...?

R [34:55-35:00]: Also es geht einfach nur um den Zusammenhang, es gibt keine Richtung.

E1 [35:00-35:02]: OK. OK.

[35:03-36:02] [brief excursus on how to read consistency matrix in EIDOS]

E1 [36:02-36:55]: Genau, also ich würde diesen Zusammenhang schon starker bewerten, also dass sich wie gesagt die Kosten für hardware minimieren, und dass dadurch das IoT-Sektor und die Anzahl der Sensorik natürlich steigt, also das würde ich für einen ganz positiven impact sehen. Wenn man die Matrix noch einmal betrachtet, wenn Network und Connection Development... Und auch das wird einen positiven impact auf den increase der Sensor haben. Weil Connectivity wird immer günstiger, momentan... ein Geschäftsmodell das Connectivity umsonst anbietet, und [inaudible] da drüber vermarktet werden, also dadurch, da sehe ich 5G und Connection Development wird natürlich der increase von Sensor stark steigen, also da sehe ich einen sehr starken Zusammenhang. Und da sehe ich eben... Connectivity wird immer günstiger und natürlich auch die neuen Standards wie 5G, wie du gesagt hast.

R [36:55-37:04]: Ja, OK. Und bezüglich der hardware-Kosten und des Netzwerks, würdest du da dann auch sagen da ist ein positiver Einfluss... äh, Zusammenhang?

E1 [37:04-37:30]: Ja. Klar. Also was ich sagte, dass du so ein Mini-Rechenzentrum mit in die Basisstation rein bringst, das kannst du nicht machen wenn die Basisstation irgendwie übermäßig teuer ist und nur einmal vorhalten muss. [brief pause]

R [37:31-37:48]: Gut. Ich glaube dann sind wir jetzt schon die ganzen fragwürdigen Felder durch. Gibt es noch irgendwelche wo du jetzt sagen würdest, da sehe ich auf jeden Fall, das passt gar nicht zusammen, oder das passt super gut? Außer das, was wir jetzt bisher schon besprochen haben?

E1 [37:49-37:50]: Jetzt allgemein? Oder?

R [37:51-37:56]: Mhm, also von den... Punkten.

E1 [37:56-38:57]: Ähm. Natürlich so ein paar Sachen wie Network und Increase of Number of Sensors, der Zusammenhang sollte klar dargestellt sein, aber es gibt natürlich jetzt, also müsste man darstellen weil euer Thema in-memory DB ist, nicht Connectivity und IoT an sich, muss man dadurch natürlich darstellen: Was ergibt sich jetzt indirekt für in-memory DB?

Dass das eben das Thema fördern wird jetzt insgesamt, also jetzt das Thema, also wenn man jetzt alles zusammen nimmt, und alles impliziert, auf das Thema in-memory DB, natürlich zu schauen hier sieht man vielleicht am Anfang nicht so die Zusammenhänge, aber klar, wenn jetzt sagen wir mal die Network, jetzt das 5G setzt sich durch, die Sensorik wird sich steigern, man möchte die real-time verarbeiten können, und braucht da eine gewisse Technologie dazu, dass dann eben sagen wir mal diesen... den Teil der in-memory DB kann dann auch... auch noch wachsen lässt. Aber ich denk das macht ihr dann ja in den Conclusions.

R [38:55-39:34]: Das ist praktisch der nächste Schritt. Da hätte ich auch noch eine Frage, und zwar, wie schätzt du das denn ein wegen der Strukturiertheit der Daten, also, könnte das nicht sein, dass die Sensoren, ja... dass das für in-memory DB schwierig wird, weil die Daten so unstrukturiert sind, und da zum Beispiel auch NoSQL-DB gewinnen? Oder kann man in-memory auf jeden Fall, je nachdem wie man die Daten dann benutzt, auch hier besonders gut anwenden?

E1 [39:34-39:27]: Ja, die Frage ist eben: Was habt ihr jetzt als in-memory definiert? Also technisch ist es klar, liegt im Speicher, ist in der Regel column-orientiert. Ich weiß nicht inwieweit ihr habt euch mit DB auseinander gesetzt habt. Aber klar, also der Punkt ist, dass du eher sowas wie eine HANA, die eher mit strukturierten Daten, das wird eher auf den betriebswirtschaftlichen Bereich gehen um Transaktionssysteme auswerten zu können. Und NoSQL hat eben, genau wie du sagst, im IoT Bereich den Schwerpunkt... ja, unstrukturierte Daten und polystrukturierte Daten zu verarbeiten, und schnell zu verarbeiten. Und wobei eben aber auch gewisse NoSQL DB auch in-memory Funktionalitäten haben, wie eben ein also die H-Base beispielsweise. Da muss man eben klar definieren was ist in-memory, und...

R [39:27-40:39]: Wir sagen relationale: Das ist die Technologie, für die wir die Szenarien machen, aber das hat natürlich... Im Rahmen dessen kann man auch die Konkurrenztechnologien dann abbilden.

E1 [40:40-42:12]: Genau, ich würde dann auch in-memory... real-time processing kann in-memory steigern das real-time processing, dieser Zusammenhang auch zwischen Network und Number of Sensors, aber insgesamt muss man da auch eben betrachten, was haben wir im Big Data Bereich... passiert. Dort eben solche Sachen NoSQL abzubilden, und interessant wird das ganze... Ich sag mal, NoSQL unstrukturiert, wenn ihr wirklich in-memory auf relational erstmal beschränkt in eurer Studie, dann ist es wie gesagt eher für die Sachen OLAP, über das irgendwie transaktionsorientierte Systeme auszuwerten, aber spannend wird das ganze Thema die Sachen irgendwann mal miteinander zu verbinden.

Das heißt, dass ich irgendwie in der Quelle unstrukturierte Daten hab, in einem Sensor, von Sensorik, und verbinde das irgendwie mit transaktionsorientierten Daten. Weiß ich jetzt nicht, was es da für use-cases gibt, aber zum Beispiel von Industrie 4.0, wenn man Produktionssteuerung hat, wo die Robotic dann irgendwie die Devices irgendwelche Daten senden, und das verbinden mit irgendwelchen Planzahlen, die aus transaktionsorientierten Systemen kommt. Und dort gibt es eben Ansätze, beispielsweise hat die Microsoft dort mit Poly-Base eine Lösung, was unstrukturierte Daten und strukturierte Daten zusammen analysieren kann. Vielleicht kannst du es ja mal kurz googeln, dann...

R [42:12-42:20]: Ich hab leider keine Verbindung. Das normale Telekom-Netz, da kann ich leider nicht...

E1 [42:20-43:18]: OK, ja. Genau, also die Oracle hat da auch sowas ähnliches im Einsatz, ich glaub das nennt sich Big Data SQL, musst du nochmal gucken bei Oracle und Microsoft. Also einmal klar, das würde ich in der Conclusion auch so mit rein nehmen, auch der Zusammenhang hier. Aber auch darüber hinaus zu schauen, wie verwendet man jetzt beide Welten – NoSQL mit RIMDB. Und dort gibt es eben beispielsweise wie Microsoft mit Poly-Base anbietet, oder auch eine Oracle... Und das kann wiederum, wenn man beides zusammen bringen möchte, wieder auch die in-memory DB wieder steigern, weil die dann natürlich wieder ein Bestandteil sind der Gesamtlösung.

R [43:19-43:31]: OK. Cool. Dann haben wir es auch schon! [inaudible]

E-Mail E4

Hallo

Es hat zwar etwas gedauert, aber jetzt bin ich die Matrix endlich durchgegangen. Hier meine Notizen dazu:

Check auf der Ergebnisebene

Demand real time response <=> RIMDB Cost

Zusammenhang erscheint mir logisch. Die Anforderungen können durch Kosten gebremst oder auch beschleunigt werden.

Umgekehrt können günstige RIMDB Kosten Bedürfnisse wecken, bzw. das Umgekehrte kann passieren.

Demand real time response <=> Development of RIMDB Open Source Solution

Auch hier ist eine Interaktion/Zusammenhang für mich nachvollziehbar. Ein breiter Bedarf kann zu einer Entwicklung innerhalb einer Open Source Community führen. Umgekehrt kann die Verfügbarkeit von (günstigen) Open Source Produkten wieder Bedarf wecken.

Demand real time response <=> Government Regulations

Hier sehe ich auch einen Zusammenhang, wenn auch etwas schwächer ausgeprägt. Durch den Bedarf (und insbesondere die Umsetzung) entsteht oft ein Handlungsdruck auf der regulatorischen Ebene. Umgekehrt beeinflusst (dann) die regulatorische Ebene die Gestaltungsmöglichkeiten im Bedarf.

Demand real time response <=> Network and Connection development

Hier hätte ich einen größeren Zusammenhang erwartet. Die Real time response von Daten bedingt auch einen entsprechend hohen Datentransfer. Hier ist die Netzwerktechnologie aus meiner Sicht ein Enabler.

Demand real time response <=> Increase in number of sensors

Hier sehe ich den Zusammenhang relativ schwach. Die Sensoren machen es eigentlich nicht. In den Projekten kommt dann IoT (internet of things) vor. Das hat natürlich einen stärkeren bis starken Zusammenhang, da hier neben den Sensoren der Datenaustausch in Vordergrund steht. (Vielleicht ist hier der Key Driver eher das internet of things, also die Sensoren und der Austausch von Daten zwischen Sensoren)

RIMDB Cost <=> Development of RIMDB Open Source Solution

Dieser Nicht-Zusammenhang hat mich im ersten Moment gewundert, jedoch war es mir relativ logisch nachdem ich mir die Projektionen noch einmal genauer angeschaut habe. Die Projektionen im RIMDB cost key driver sind alle sehr Hardware lastig. Dies hat in der Tat keinen Zusammenhang mit der Open Source Entwicklung, die ja einen Software Fokus hat.

RIMDB Cost <=> Government Regulations

Hier gibt es solange keine Zusammenhänge, bis regulatorische Anforderungen Kosten verursachen, oder durch staatl. Subventionen Kosten gesenkt werden.

Ein Zusammenhang ist mir hier an folgender Stelle nicht klar.

RIMDB Cost <=> Network and Connection development

Aus meiner Sicht sind diese beiden Key Driver unabhängig voneinander. Hier sehe ich auch wenig Beeinflussung. Den Zusammenhang zwischen der Cloud Projektion und der Netzwerk Entwicklung ist für mich schlüssig.

RIMDB Cost <=> Increase in number of sensors

Hier sehe ich den Zusammenhang nicht so. Beides sind für mich eher unabhängige Teile. Z.B. Reduction of hardware cost und IoT is largely in place erschließt sich mir nicht.

Development of RIMDB Open Source Solution <=> Government Regulations

Hier sehe ich auch nur wenig Zusammenhang. (Es sei denn es gibt eine Initiative der öffentlichen Verwaltung verstärkt auf Open Source Lösungen zu setzen wie es in der Vergangenheit schon war. Dies ist aus meiner Sicht aber nicht absehbar.)

Development of RIMDB Open Source Solution $\Leftrightarrow$ Network and Connection development

Ich sehe da auch keinen Zusammenhang

Development of RIMDB Open Source Solution $\Leftrightarrow$ Increase in number of sensors

Ich sehe da auch keinen Zusammenhang

Government Regulations $\Leftrightarrow$ Network and Connection development

An einer Stelle würde ich stärkeren Zusammenhang sehen: Government Regulations mit der Projektion Development jump through government investment.

Government Regulations $\Leftrightarrow$ Increase in number of sensors

Hier sehe ich auch einen starken Zusammenhang, da das IoT von Regulierung betroffen sein wird und andererseits (siehe aktuell Digitale Strategie 2025 der Bundesregierung) wird das IoT staatl. gefördert werden.

Network and Connection development $\Leftrightarrow$ Increase in number of sensors

Das bunte Bild konnte ich soweit nachvollziehen. Der generelle Zusammenhang den ich hier sehe ist, dass das IoT zur Kommunikation ein starkes Netzwerk benötigt.

Viele grüße
Jörg Frank

E-Mail E6

Not available, please contact the researcher for more information

Attachment D: Interview Transcript Phase R4

Transcript E10

Date: 25.02.2016

Duration: 35:55 [mm:ss]

Modus: face-to-face interview

Interviewee:

Expert 10 [E10]: Lars Geißler; Founder and CEO of WebXells GmbH in Potsdam

Interviewers:

Julia Reichelt, Researcher [R]

Victoria Götz, Thesis Supervisor [TS]

E10 [00:00-00:03]: Gut, heute die einzelnen Szenarios, wie viele haben wir, vier Stück?

R [00:03-00:41]: Genau, vier Stück. Eines nach dem anderen angucken, also immer auch wenn die sich überlappen, es geht ja immer um die Kombination der verschiedenen Faktoren. Und das heißt immer dann separat betrachten – OK, das ist der Stand praktisch, was bedeutet das jetzt. Auch so ein bisschen, ich muss die praktisch im Nachhinein noch ein bisschen mehr drum rum beschreiben. Wie kam es dazu, was könnten mögliche Einflussfaktoren sein, wie... Ja. Wie sieht praktisch das Feld so aus unter diesen Bedingungen. Und dann die Brücke schlagen und was bedeutet das dann wiederum für Business Operations.

E10 [00:41-00:47]: OK, dann bin ich gespannt. Fangen wir an. Cool ist, dass schon mal eine Struktur da ist. Das ist cool, weil das vergleichbar wird.

R [00:48-00:55]: OK. Genau, also vielleicht kannst du das erstmal durchlesen, das erste Szenario.

E10 [00:55-01:04]: Na stell es doch vielleicht einfach vor, lass uns doch gleich im Gespräch das gleich, also dass du jeden einzelnen Punkt mit mir durchgehst.

R [01:04-01:43]: Können wir machen. Also beim ersten geht es darum, dass die Nachfrage nach real-time Analysen hoch ist – mittel bis hoch. Und real-time ist einfach der State of the Art geworden. Das heißt es gibt noch Firmen, vor allem größere Firmen, die hinterher hinken, noch nicht real-time... Noch nicht so weit sind in der Implementierung von real-time Systemen, aber die müssen jetzt schnell nach ziehen um praktisch nicht aus dem Markt zu fallen. Das heißt weiterhin hohe Demand, auch für Systeme mit real-time Analysen.

E10 [01:44-01:56]: Weißt du, was ich mich da immer gleich frage – entschuldige bitte, wir wollen ja darüber diskutieren, oder? Oder willst du dein Schema durchgehen? Weil ansonsten würde ich immer gerne gleich Bezug nehmen. Die erste Frage die sich mir stellt ist immer: Fallen die wirklich aus dem Markt?

R [01:56-01:57]: Das ist die Frage...

E10 [01:58-03:05]: Es wird nämlich nachher wichtig, wenn du den Kreis drehen willst HANA zu verkaufen. Musst ja irgendwann über Argumente nachdenken: Was macht das denn so gefährlich, wenn du nicht in real-time arbeitest. Und das wäre die erste Frage, die ich mir immer bei so einer Sache stellen würde. Was bedeutet denn das für, ja Versicherungen – meinetwegen, wenn sie nicht in real-time ihre Preise anpassen können. Nehmen wir mal an: Wetter hat sich geändert, Glatteis ist jetzt öfter hier in der Bundesrepublik, und die Autotarife müssten angepasst werden. Andere Versicherungen, die HANA einsetzen, machen das in Echtzeit, können vielleicht auch ganz andere Produkte plötzlich den Leuten anbieten. Weil wenn du in Brandenburg bist hast du nicht so viel Eis wie unten in Bayern, und kannst dann praktisch in Echtzeit einen Autotarif anbieten. Das wäre ja ein ganz anderes Produkt: kein Festpreis-Produkt, sondern ein Produkt, das sich anpasst. Und das sind dann so die Argumente, die dann geil werden für HANA. Also deswegen nochmal, die Frage – [phone rings] entschuldigt bitte – passiert das wirklich, also fallen die aus dem Markt, oder nicht.

R [03:05-03:22]: Das ist eine berechnete Frage. Aber... hier ist es praktisch nur der Status: OK, die haben... Es ist der State of the Art, also in zehn Jahren praktisch ist es State of the Art, dass man real-time benutzt – wie auch immer es dazu gekommen ist.

E10 [03:22-03:23]: Ah OK, das setzt du also einfach fest.

R [03:23-03:24]: Genau.

E10 [03:24-03:25]: Gut.

R [03:25-05:19]: Und es geht halt so ein bisschen auch in die Kombination der Faktoren, das heißt wenn wir die jetzt alle einzeln angucken, dann wiederholt es sich ganz viel. Es geht aber mehr so ein bisschen darum auch anzugucken: OK, was bedeutet das in Verbindung. Das heißt wir müssen die natürlich schon auch einmal durchgehen, aber dann insgesamt angucken.

OK, lassen wir das erstmal als ersten Punkt stehen. Dann als zweiten Punkt haben wir die Kosten. Da – also für RIMDB – an sich, die sind stagniert. Und als Hauptgrund dafür hier, liegt daran, dass die Hardware zwar an sich günstiger geworden ist, aber dafür die Datenvolumen so gestiegen sind, dass die Effizienz auch gesteigert werden muss. Das heißt im Grunde hat sich nicht so viel geändert für den Endkunden.

Dann als nächstes haben wir eine Entwicklung, dass open-source Datenbanken [DB] den Markt übernommen haben, das heißt die kommerziellen Anbieter wurden verdrängt. Und es werden jetzt RIMDB als open-source angeboten.

Dann als nächstes sieht es so aus, dass von Regierungsseite und von den Regularien aus, die... praktisch eine Offenheit der Daten vorangetrieben wird, also es gibt keine besonders starken Restriktionen, was Privacy, Security, diese Themen angeht. Und die Regierungen versuchen da aber möglichst großen Nutzen draus zu ziehen. Das heißt zum Beispiel auch – Stichwort Industriespionage, also die versuchen da wirklich auch selber, dadurch, dass sie es so offen lassen, auch selbst ihre Infos raus zu ziehen. [E10's phone starts vibrating]

Und dann als nächstes – bin ich das – ne, OK.

E10 [05:19-05:22]: Sorry, das ist mein normales Leben.

R [05:22-06:53]: Und dann als nächstes haben wir das Netzwerk als großen Einflussfaktor, und da gab es dann die Entwicklung, dass sich das Netzwerk und die Verbindung besser ausgeweitet hat, das heißt jetzt könnte man es noch als bottleneck für das Internet of Things [IoT] und diese ganzen Sensoren sehen, weil es eben noch nicht flächendeckend und schnell genug ist. In diesem Szenario ist allerdings das Netzwerk schon... ja, gut dabei, und ist praktisch kein Hinderungsgrund mehr. Und das kam vor allem dadurch, dass auch aus der Industrie und privaten Investment am Netzwerk weiter... Oder sagen wir mal an der Connectivity weiter gearbeitet wurde. Das heißt es gibt... Es gab nicht nur die Telekom als Netzwerk... die das Netzwerk bauen, sondern es gab auch noch andere Lösungen, die da dagegen gehalten haben, wodurch dann auch so eine Art Konkurrenz entstand, und insgesamt die Entwicklung voran getrieben wurde.

Und als allerletztes: Die Sensoren und das IoT haben sich tatsächlich so ausgebreitet, wie es vorher gesagt wird. Das heißt diese ganzen Prognosen sind praktisch eingetreten.

Das wäre jetzt praktisch der Stand. Und die Frage ist: Was bedeutet das dann für RIMDB als erstes, und im zweiten Schritt für Operations.

E10 [06:54-08:28]: Naja, die Frage, die kannst du ja mit dem Szenario ja schon selber beantworten, also das ist ja klar. Also wenn das wirklich so eintritt, wie du das jetzt in dem Szenario A geschildert hast, dann kommt man darum ja gar nicht mehr drum rum – es sei denn, es gibt eine andere Technologie, die genauso schnell, oder schneller zu Verfügung steht. Da gibt es ja auch verschiedene Möglichkeiten, oder verschiedene Technologien, die auf dem Markt sind, die sich gerade entwickeln, um sowas auch abzubilden.

Und für Operations, das hatten wir ja schon zum Anfang. Also weil sich jetzt plötzlich nämlich Services und Produkte komplett verändern können. Du kannst, wenn wir beim Beispiel der Versicherung bleiben, kannst du plötzlich Produkte ganz anders aufbauen. Nämlich auf Echtzeit-basierten Daten. Und wenn wir das jetzt noch weiter spinnen in unserem Beispiel, wenn da jetzt Connectivity im Car-Umfeld nimmst, und die Autos vernetzt sind, und die Sensorik vom Auto permanent Informationen und Daten einspielt, dann könntest du sogar Tarife bilden bei einer Versicherung, die auch in Echtzeit funktionieren. Ohne menschliche Entscheidung. Dann wird es interessant. Weil es immer darum geht bei solchen Technologien, dass man eben den Menschen außen vor lässt.

Also zumindest für mich in meiner Denke. Und dann hätte man für den [inaudible]-Bereiche eine massive Veränderung, wenn Produkte neu denken, Services neu denken.

R [08:29-08:42]: Und wie würdest du das jetzt persönlich einschätzen – haben RIMDB da auch die Chance, dass die da wirklich die Technologie werden, mit der das alles ausgeführt wird. Oder siehst du da eher andere Technologien?

E10 [08:43-09:33]: Teilmärkte! Es kommt immer ganz darauf an, was du machen willst. Wenn du Video-Streams auswerten willst würde ich zum Beispiel eine andere Technologie nehmen. Ich selber, WebXells, wir arbeiten ziemlich viel im 3D Umfeld. Wenn du einen 3D-Stream hast, dann irgendetwas von Räumen, Gebäuden, oder oder oder, würde ich auch andere Technologien nehmen. Wenn du reine Sensorwerte nimmst, dann macht das durchaus Sinn. Also wenn wir bei unserem Beispiel bleiben, dann würde ich sagen: Jo, da wird sich HANA mit Sicherheit durchsetzen. Also HANA wird in dem Markt mit Sicherheit einen Platz haben. Das glaube ich schon. Aber soweit ich weiß entwickelt SAP ja selber schon an Nachfolgeprodukten. Oder? Also kommt... Also ist der nächste Step auch da schon geplant.

R [09:34-09:44]: OK. Gut, das heißt du siehst dann auch diese real-time service delivery, oder Anpassung von Tarifen als zentrales...

E10 [09:42-10:30]: Ganz klar, ganz klar. Es wird ein wirklicher Brecher werden für viele Produkte. Weil Firmen, die das ja dann nicht machen, wie ja unser Beispiel: Versicherungen, die das nicht machen, ganz klar, die fallen weg. Und die Versicherung, die es als erstes macht, dir nämlich die Möglichkeit gibt, weil du, hey, in Brandenburg lebst, ohne drei Meter Schnee, du wohnst nicht irgendwo auf einem Berg, und hast dadurch einen Kostenvorteil von 200-300 Euro, dann wirst du diese Versicherung nehmen. Das ist für dich überhaupt keine Frage. Und für dich wird auch keine Frage sein, die Erlaubnis, oder das Häkchen zu setzen, dass dein Auto – wenn du 300 Euro im Jahr sparst – deine Daten [inaudible] weiter leitet. Das interessiert dich überhaupt gar nicht. Also ja: Ich glaube es hat eine Berechtigung, und ja: Es wird verdammt viel verändern.

R [10:31-10:39]: Und könntest du dir Bereiche vorstellen, die dadurch irgendwie wegfallen oder ein massives Problem bekommen, weil sie das nicht so implementieren können?

E10 [10:39-11:08]: Weil sie es nicht implementieren können...? Ne, also das glaube ich gar nicht. Alle Bereiche... mal kurz drüber nachdenken... Alle Bereiche, die ich mir jetzt so mal vorstellen kann, können digitalisiert werden, und zwar so weit, dass du den Menschen außen vor lassen kannst. Also ich wüsste jetzt keinen Bereich, selbst in der menschlichen Pflege, wo der Mensch nicht ersetzbar wäre. Oder was hast du gesagt, 2025, oder 2020...

R [11:08-11:10]: Genau, 2025. Also zehn Jahre, neun Jahre.

E10 [11:10-11:36]: Also in zehn Jahren wird die Verschmelzung zwischen Mensch und Maschine noch nicht so weit sein, also wesentlich weiter als jetzt zwar, aber noch nicht so weit sein, dass man von einer echten Verschmelzung reden kann. Aber auch da glaube ich geht es auf jeden Fall in die Richtung. Und auch da sind viele schnelle Datenbanklösungen, um Entscheidungen zu treffen auf Grundlage von Daten, unheimlich wichtig.

R [11:37-12:03]: Gut. Noch zum Thema hier... Wir haben ja den einen Punkt auch, dass die Regierungen da auch versuchen viel Nutzen draus zu ziehen. Was glaubst du könnten da zum Beispiel Auswirkungen sein? Also was passiert mit diesen Daten, welche Daten brauchen die überhaupt wenn die jetzt zum Beispiel so ein Unternehmen mit überwachen...

E10 [12:01-13:49]: Keine Grenzen... Es gibt keine Grenzen. Schau mal, jeder in einem Staatsapparat, auch jeder in einer Firma, das ist überall das gleiche, jeder muss auf seiner Stelle beweisen, jeden Tag, dass er das Geld, was er bekommt, zurecht bekommt, nämlich, dass er seine Arbeit tut, das muss er beweisen. Auswertungen, Berichte, sind genau dafür da, zu sagen: Hey! Ich tue meinen Job! Alles was die Maschine zuliefern kann um genau das zu tun wird mit Kussband genommen. Das größte Problem, meiner Erfahrung, bei der Einführung bei größeren Software-Projekten, oder Technik-Projekten, ist immer der Mensch. Die Menschen, die damit arbeiten müssen.

Wenn du den Menschen nicht verkaufen kannst, dass er einen Mehr... oder einen Zugewinn davon hat von der Einführung der Software oder von der Technologie, dann wird er sich dagegen stellen, weil er Angst vor Veränderungen hat. Vollkommen normal. Wenn du einem Regierungsbeamten sagst: Hey! Das nächste Mal, wenn dein Parlamentarier dich fragt: Wie viele syrische Flüchtlinge gibt es gerade in Brandenburg, die eine

Ausbildung als Meister haben, eine syrische, wenn der das auf Knopfdruck kriegt, dann wird er sofort sagen: Jo! Die Software will ich. Das ist das größte Problem, zum Beispiel im Regierungsapparat, kriege irgendeine Anfrage, irgendein Schwachsinn, die müssen jetzt anfangen zu recherchieren: Wo bringe ich die Daten her. Und wenn die Daten verfügbar wären – und die sind ja eigentlich da, bloß keiner weiß wo, und wie sie daran kommen, dann ist man auf dem richtigen Weg. Also wird massive Auswirkungen haben, so eine Technologie.

Und eigentlich verstehe ich gar nicht warum es da Probleme gibt. Ja doch, ich verstehe schon, warum es da Probleme gibt.

R [13:49-14:10]: OK. Und dann vielleicht noch als letztes: diese Entwicklung von parallel Netzwerk, neue Netzwerk-Operator, neue Netzwerk-Anbieter, oder auch neue Möglichkeiten von Connectivity. Was fällt dir dazu vielleicht noch ein?

E10 [14:10-14:26]: Boah. Schwer. Damit beschäftige ich mich weniger. Also da kann ich keine Aussage mehr mit treffen. Also was das dann technologisch bedeutet, also im nächsten Schritt – weiß ich nicht. Da habe ich mir auch noch keine Gedanken drüber gemacht.

R [14:26-15:32]: OK, cool. Dann können wir das gleich schon mal so stehen lassen würde ich sagen, und dann zum nächsten gehen. Da habe ich auch schon mal markiert was jetzt die drei Sachen sind, die anders sind. Und die noch nicht vorher gehört wurden.

Das heißt: Einmal gibt es eine Veränderung bei den Kosten. Das heißt die Kosten für RIMDB sind massiv gesunken, dadurch, dass man immer mehr Cloud-Systeme benutzt, und nicht so auf die hardware praktisch mehr ankommt. Dann gibt es sowohl Open-Source als auch kommerzielle Lösungen. Und dann von Regierungsseite ist es eigentlich relativ ähnlich, bloß, dass noch viel mehr unterstützt wird aktiv, um praktisch Mehrwert, auch wirtschaftlich gesehen, aus diesen Daten zu nutzen. Dass man zum Beispiel auch die Workforce in die Richtung ausbildet, dass genug Leute da sind, die auch das damit machen können, oder Forschung unterstützt, und so weiter.

E10 [15:32-17:28]: Was ich als unrealistisch empfinde, ist, dass die Kosten fallen auf einer DB-Seite, oder Service-Seite. Das wird nicht passieren. Und warum? Die Daten werden mehr. Die werden massiv mehr. Schau, wie sich die letzten drei Jahre entwickelt haben. [inaudible], was jetzt auch schon Daten auch mitnimmt. Ich kenne ganz viele die hier mit den Bones rum rennen, hier mit diesen Fitnessarmbändern, allein wenn du dir da mal das Datenvolumen rein ziehst, was die an Daten aufnehmen, als rein checken, das ist immens. Also die Kosten werden gleich bleiben. Und ich glaub auch nicht prozentual gleich bleiben, ich glaube nicht, dass es zu einer Minderung kommt, bloß weil da eine Cloud genutzt wird, das ist unrealistisch.

Ich glaub, die Regierungen werden Interesse daran haben, gerade in der jetzigen Entwicklung. Ich selber bin einer Arbeitsgruppe drin wo es um eine technische Lösung für die Flüchtlingsproblematik geht. Und da geht es sehr sehr [inaudible]. Nur, dass 20,000 Stellen irgendwelche Daten erfassen, und dass man guckt, dass man data migration service baut, letztlich, der die Daten zusammenfasst, um die dann in verschiedenen Kanälen zu veröffentlichen. Klar, da ist definitiv ein Interesse da. Und die werden das nicht blockieren.

Das einzige, was blockieren kann – noch – sind private Interessengemeinschaften. Es wird immer Gegenbewegungen geben. Immer, wenn du so eine starke Veränderung hast wie – Industrie 4.0 finde ich doof – wie das Zusammenwachsen Mensch-Maschine, das ist für mich nämlich das eigentliche Thema, wird es eine starke Gegenbewegung geben. Die Gegenbewegung ist aber aus meiner Sicht so schwach, dass sie das nicht verhindern kann, dass diese Veränderung kommt.

R [17:28-17:30]: Gegenbewegung eher aus ethischer Sicht dann...

E10 [17:30-18:11]: Ja, ich denke schon. Hast du ja jetzt auch schon. Also es gibt Leute, die sich gegen Social Communities ganz massiv stellen, und es gibt Leute die es halt mit nutzen, aber die Mehrheit überwiegt erstmal das mit zu nutzen. Im Moment kippt das gerade, das ist sehr interessant, auch für mich. Also was in den Zahlen ist, dass immer mehr Leute aus den Social Communities raus gehen – aber andere Dienste dafür umso mehr nutzen. Das ist eine hochinteressante Geschichte.

Aber da geht es um Arbeit und Geld verdienen. Also Arbeit wird sich verändern. Glaube ich nicht, dass einer von uns noch jemals irgendwo in der Produktion arbeitet als Facharbeiter.

R [18:12-18:17]: So ganz global heißt das ja dann, dass schon Arbeitsplätze weg fallen, oder glaubst du die werden ersetzt?

E10 [18:17-19:20]: Hat man ja jetzt schon. Das sieht man jetzt schon. Da brauchst du ja nur in die Geschichte gucken. Bei jeder industriellen Revolution, und meiner Meinung nach sind wir gerade in einer, sind erstmal komplett die ganzen alten Arbeitsplätze gecuttet worden. Und dann irgendwann haben sich neue aufgebracht. Ich glaube nicht, dass irgendjemand in Zukunft am Fließband stehen wird. Warum?! Du kannst Kapital einsetzen, und das Kapital als Produktionsmedium reicht aus, um alles herzustellen, was du willst. Geistige Arbeit, kreative Arbeit, selbstständige Arbeit, Arbeit im Dienstleistungsbereich: ja. Dieses Jobsegment steht.

Da ist aber Politik, oder Gesellschaft, gefragt, ne. Diese Methodiken und Werkzeuge innerhalb einer Gesellschaft, wie die Gesellschaft funktioniert, anpassen. Kostenloses Grundeinkommen, und und und, gibt es ja tausend Möglichkeiten, wie man da als Gesellschaft drauf reagieren kann. Was auch OK ist, um die Leute auch wieder aufzufangen. Ich persönlich glaube, dass wir im nächsten Jahr 2-3 Millionen Arbeitsplätze hier in Deutschland verlieren werden. Ganz persönlich.

R [19:21-19:37]: Das heißt das könnte hier auch sehr gut mit rein spielen, diesen Bereich, also OK, die Regierung unterstützt vielleicht diese Bewegung Richtung IoT, Big Data, Mensch-Maschine, aber sie müssen andererseits auch das dann auf einer anderen Ebene wieder auffangen.

E10 [19:26-21:06]: Ja sozioökonomisch hat das Folgen das Ganze. Man muss halt rechtzeitig vieles anpassen. Die Bildung, das ist ein ganz wesentlicher Punkt. Duale Ausbildung: Fragezeichen Notwendig? Keine Ahnung. Aber müsste man eigentlich, wenn du das Land, die Weichen vernünftig stellen willst, müsste man sich jetzt eigentlich darüber Gedanken machen und genau diese Dinge anpacken. Wird glaube ich nur zum Teil gemacht, weil die Leute gar nicht die Probleme im Moment sehen. Die sind glaube ich...

Also ich bin im DIHK auch noch mit drin, das ist so die Deutschland-Vereinigung von den ganzen IHKs, und dort gibt es auch eine Industrie 4.0 Buzzword-Community, und ich bin da in einer Arbeitsgruppe, die daran arbeiten. Und die sozioökonomischen Folgen, die spielen so erst gaaaanz ganz langsam so wirklich eine Rolle, aber wirklich eher auf Gesetzgebungsebene, also wie muss ich jetzt Recht anpassen, oder so, aber noch nicht so im Thema Arbeitsmarktpolitik, Bildungspolitik und und und. Das kommt aber, da bin ich mir sicher. Aber keine Regierung kommt drum rum, das zu tun. Also zumindest nicht in den westlichen Ländern, ansonsten dass zu diesem Step nicht schaffen. Also das, grad China, Westen, USA, Russland – jetzt hätte ich beinahe gesagt UdSSR, dann musst du als Industrienation, gerade Deutschland, musst du dich irgendwie positionieren. Und das ist wirklich praktisch **inaudible** zu tun.

R [21:06-21:25]: Gut, nochmal zu dem einen Punkt hier: sowohl Open-Source als auch kommerzielle RIMDB. Jetzt innerhalb von diesem Setting. Wer glaubst du nutzt was? Wo spaltet sich das, also wo sind die Markträume?

E10 [21:23-21:27]: Wie ist das jetzt gerade?... Wie ist das denn jetzt gerade?

R [21:28-21:32]: Ich würde sagen, große Firmen nehmen kommerzielle Lösungen und...

E10 [21:30-21:32]: Genau, und warum?

R [21:32-21:34]: Weil sie es sich leisten können...

E10 [21:33-22:38]: Nee! Weil es nachhaltig ist! Die haben jemanden, dem sie auf den Kopf hauen können, wenn es nicht funktioniert. Ganz einfaches Ding. Du hast einen Ansprechpartner. Was interessiert mich das Geld? Das Geld ist vollkommen egal. Geld... Musst halt bloß gucken, dass deine Liquidität passt, und dass deine Anteilseigner zufrieden sind, aber Geld an sich spielt überhaupt gar keine Rolle. Wichtig ist, dass du nachhaltig wirtschaften kannst. Und dass das Sinn macht, was du tust. Wenn du das Ding in eine Open-Source DB rein haust, dann brauchst du Personal. Und was Nahe liegt: Menschen sind sehr sehr schwer zu beherrschen. Eine andere Firma, die seit – was weiß ich – zehn Jahren am Markt ist, da weiß ich: OK, da hab ich Hebel und Werkzeuge, ich kann die **inaudible**. Aber einen einzelnen Menschen...

Kleinere Firmen, mittlere Firmen werden Open-Source nehmen. Große Industrieunternehmen werden... Konzerne, die international arbeiten, werden definitiv andere Firmen nehmen, bezahlte Lösungen. Also so wie jetzt. Genau das gleiche.

R [22:38-23:48]: Gut. Ich glaube wir haben alle Punkte in diesem besprochen. Dann zum nächsten, hier auch wieder markiert, die die vorher noch nicht vorkamen praktisch. Da geht es dann einmal darum... Da wäre... Genau, jetzt auf Open-Source wieder bezogen, das Szenario, dass die großen Firmen, wie zum Beispiel eine SAP, oder IBM, ihre Open-Source, äh, ihre RIMDB, Open-Source anbieten als Vehikel für andere Lösungen. Das heißt der Fokus geht so ein bisschen weg von RIMDB, und sie benutzen das nur um andere Systeme, zum Beispiel in verschiedenen Bereichen zu etablieren, machbar zu machen, genau. Und das zweite ist, dass das Netzwerk sich zwar auch verbessert hat, aber eher von einem großen Regierungsinvestment kommt, im Gegensatz zu diesem Split von Private und Regierung.

E10 [23:47-23:48]: Das habe ich nicht verstanden.

R [23:49-24:13]: Das heißt: In den vorherigen Szenarien war es so, dass Connectivity an sich sich praktisch verbreitet hat, weil es verschieden private Investments da rein gab, um das auszubreiten, und jetzt war es praktisch ein großer Regierungsaufwand, der da massiv investiert hat, um schnelleres und mehr Connectivity zu erstellen.

E10 [24:13-25:05]: Also, dann wieder unten angefangen: Das glaube ich nicht. Sondern es wird immer in Mix sein aus privaten Investitionen, und aus politisch getriebenen Motiven, und da wird auch Investitionen, und... Das... Die Antwort oben... Ist schwer zu sagen... Schwer zu sagen... [brief thinking pause] Das kann ich dir nicht beantworten.

R [25:05-25:10]: Und jetzt hier oben auf die bezogen? Auf die Systeme?

E10 [25:09-25:18]: Mhm, genau, genau. Also... Nee, kann ich dir nicht beantworten leider.

R [25:18-25:49]: Zum Beispiel, also was ich mir jetzt so überlegt hatte war, dass eben es andere, gerade in der Kopplung... Im Bereich Produktion sagen wir jetzt mal, dass zum Beispiel RIMDB eher Daten zuliefern, aus den Planungsprozessen zum Beispiel, und dann die Sensordaten, weil die halt auch doch schneller bearbeitet werden, oder anders, dass es keine strukturierten Daten sind, dann eben...

E10 [25:47-27:38]: Ich würde es noch anders angehen. Weißt du, wenn du – nehmen wir mal jetzt am Beispiel HANA und SAP. Wenn du jetzt zu jemandem hin gehst und willst HANA verkaufen, sagen: Hier habt ihr eine unheimlich schnelle Datenlösung. Dann werden sie dir sagen: OK, wofür? Und dann sind wir wieder beim Punkt: Es geht eigentlich den Leuten nicht um die Technologie, die eingesetzt wird, das ist vollkommen egal, sondern es geht den Leuten um die Use Case. Und den Gewinn, den sie aus diesem Use Case rausziehen. Gar nicht monetär gemeint, sondern es kann Zeitersparnis sein, kann mehr Daten sein, kann keine Ahnung was. Das heißt: Wenn du wieder aus der Verkaufsdenke raus gehst, dann wäre es geschickter, wenn man HANA Open-Source macht, komplett Open-Source, und einen Framework anbietest, wo du Use Cases drauf schreiben kannst, um HANA zu benutzen. Um diese Geschwindigkeitskomponente zu haben. Und SAP verkauft einzelne Module, die du für irgendwas brauchst. Das würde mehr Sinn machen für so eine Technologie.

Ich glaub so kriegst du auch Abhängigkeiten – Technologie-Abhängigkeiten – besser verpackt, und damit auch besser verkauft. Weil keiner will sich eigentlich heutzutage nur noch an eine Technologie binden, das ist immer problematisch, da musst du schon echt mehr Erleben haben, da musst du schon echt Mehrwert haben auch. Und wenn du jetzt plötzlich ein Framework hättest, wo du selber dir Lösungen schreiben kannst für HANA, dann ist das so ein Mehrwert. Und dann kaufst du auch mal den Service XY mit dazu, weil jetzt kriegst du noch Fluggastdaten meinetwegen, oder du kriegst noch die Daten. Und du kriegst noch die Lösung, die Berechnung. Weißt du was ich meine? Also so würde es für mich eher einen Sinn machen.

R [27:39-28:13]: OK, cool. Gut. Jetzt hätte ich hier noch... Investment... Ja, das andere ist eigentlich relativ ähnlich zu den Sachen, die wir schon vorher besprochen haben. Dann: Vielleicht können wir nochmal kurz überlegen, hier gibt es da irgendwelche anderen Einflüsse auf den Bereich Produktion, Operations? Kann sich dadurch irgendwas verändern?

E10 [28:13-29:36]: Na, dass die Leute, also wenn es so passieren würde wie ich das eben gerade gesagt habe zum Schluss, dann würdest du plötzlich Firmen, Anwender, Entwickler mit rein nehmen, die irgendwelche Lösungen bauen, die vorher gar nicht befähigt waren dazu, ohne [inaudible] einzusetzen. Du multiplizierst praktisch das Ganze, auch in der Geschwindigkeit. Also guck dir Apple an, das Framework X-Code ist komplett free, jeder kann es nutzen, jeder kann Anwendungen drauf schreiben. Apple verdient erst dann da dran,

deswegen ist es so genial, dieses Geschäftsmodell, wenn ein Entwickler etwas verkauft. Davor hat der keine Kosten – ja stimmt nicht ganz, \$100 kostet glaube ich der Zugriff, aber hey! Das kann sich jeder plötzlich leisten! Und jeder kommt dann plötzlich in diesen Goldtausch mit rein. Kann SAP auch machen mit HANA, oder solchen Technologien, wenn du dir plötzlich nämlich was ganz anderes überlegst. Nicht: Wie kriege ich irgendeine Technologie verkauft, sondern wie schaffe ich einen Marktplatz von Menschen, wie motiviere ich diese Menschen, diese Technologie zu nutzen. Welche Use Cases kann ich denen schon mal schenken, kann ich denen vorgeben, vordenken, sodass die Leute aktiviert werden selber da rein zu gehen und [inaudible].

R [29:36-30:18]: OK. Cool. Gut, dann können wir schon zum letzten gehen. Da ist gar nicht so viel anders, beziehungsweise, die Faktoren haben wir alle schon mal gehört, aber die sind jetzt ein bisschen anders kombiniert, und das, was vorher noch nicht vorkam, ist dass – ja, es eigentlich so eine richtige Marktsättigung gibt an diesen real-time Analysen, also jeder hat sich irgend so ein System installiert, und man braucht vielleicht noch Updates, aber es gibt jetzt nicht... Die ganzen response-times werden jetzt nicht so massiv schneller, dass man da noch größeren Mehrwert an neueren Systemen haben könnte.

E10 [30:16-33:27]: [inaudible] Wird klar [inaudible]. Wie mit jeder Technologie hast du halt irgendwann eine Marktsättigung, und dann kommt das nächste, kommt die nächste Technologie. Es geht halt nicht immer um Auswertungen, sondern es geht echt immer um die Use Cases. Welche Produkte und Services kannst du damit anbieten, das ist für mich eigentlich das einzig interessante. Tut mir leid, ich bin schrecklich, weil ich immer nur an Kohle verdienen denke. Ich denk immer nur: Wie kann man etwas verkaufen. Irgendwann ist der Markt gesättigt, da kommst du nicht drum rum. Du musst halt gucken in dem Lebenszyklus: Wann fang ich an, den nächsten Step zu denken? Das kannst du jetzt noch nicht, weil diese Technologie noch gar nicht im Markt ist. Du weißt nicht welche Veränderungen das eigentlich mit sich bringt.

Für Menschen – weiß ich nicht, vielleicht wird es irgendwann möglich sein, Prothesen mit HANA bauen. Hört sich total crazy an, aber wäre doch geil, wenn deine Sensorik am Körper merken würde, in welcher Situation du dich gerade befindest, und dann halt dementsprechend deine Prothese vorher schon reagiert, bevor du es eigentlich denkst. Das – du lachst da drüber, aber wenn ihr euch... Amerika, du bist ja bald im Heiligtum, in Stanford, wenn ihr euch da die Robotik anguckt, und Medizin-Robotik, und woran die arbeiten – dann ist das genau das: Verschmelzung von Technologien und der Biogenetik, und dem, was technisch möglich ist.

Und dann kannst du plötzlich wieder ganz andere Anwendungsfälle wieder denken für HANA. Dann geht es nicht um Echtzeit-Auswertungen, sondern dann geht es plötzlich um Entschlussfassung: Also wie kannst du Entschlussfassung, also wer wie was wann wo, beeinflussen. Vielleicht kannst du Entschlüsse treffen, bevor sie getroffen werden müssen. Bevor du, als Mensch, daran denkst. Und dann wird's geil. Dann wird's echt, also, dann wird's von der Denke auch, also bei mir, zumindest im Moment, auch verschoben, weil da bin ich leider nicht intelligent genug, weil dann fängst du nämlich an im Raum über [inaudible] Dimensionen zu sprechen. Ja, tut mir leid, da steige ich aus.

Aber das sind so Ansätze, in die man rein gehen kann, die man jetzt aber noch nicht sieht, weil jetzt ist das eigentliche Problem noch gar nicht gelöst. Ich glaube wenn jetzt ein Vertriebler von SAP zu einer Versicherung hin geht und sagt da: Hey, wir haben eine coole neue Technologie! Ich hab selber mal für SAP gearbeitet, und weiß, dass die Vertriebler keine Ahnung von der Technik haben, und die können gar nicht erklären, welche Use Cases jetzt aufgemacht werden können, und was die Versicherung jetzt für neue Produkte anbieten kann. Die denken sich jetzt leider meist nicht in den Kunden rein. Das ist nämlich das Problem. Deswegen muss jetzt erstmal das eine Problem gelöst werden, und dann kannst du hier ins nächste Szenario rein gehen und überlegen, was bedeutet das jetzt für die nächste Art der Technologie.

Was braucht man eigentlich alles für Technologie? HANA ist jetzt eine Technologie, zumindest von meinem Empfinden, die gemacht wurde weil sie möglich ist. Aber es wurde noch nicht drüber nachgedacht, was ist eigentlich damit möglich. Also: Die falsche Frage wurde gestellt. Und da kommt es immer drauf an glaube ich.

R [33:28-33:36]: Und hier in diesem Szenario wäre es auch wieder, dass Open-Source wieder den Markt übernehmen, aber das passt ja dann eigentlich auch schon ganz gut zu dem...

E10 [33:36-33:46]: Aber ich glaub nicht daran, dass eine Seite die andere Seite übernimmt. Die Berechtigung ist für beide da. Also ich wüsste nicht aus welchem Grund sollte das passieren.

R [33:47-34:31]: Was waren denn da die... die... Begründungen. Ja, also genau: Wie wir darauf gekommen waren, war auch von vorherigen Interviews, dass eben es so eine ähnliche Entwicklung gab in anderen Bereichen von

Software. Dass dann Stück für Stück die Open-Source Lösungen einfach immer besser wurden, und dann die kommerziellen praktisch in einen Zugzwang kommen, und sich dann praktisch eher etwas etabliert, dass Open-Source der Standard ist, und dann schon noch andere Services außen rum angeboten werden, aber nicht so jemand wie HANA so ein...

E10 [34:29-35:35]: Aber, also, das... Ich kenne mich bei Suchmaschinen zum Beispiel nur aus. Bei Suchmaschinen hast du vollkommen Recht. Es gibt keine vernünftig kommerziellen Suchmaschinen. SAP hatte T-Rex oder irgendwie sowas mal, Google hatte ihre Lösung, das hat alles nicht richtig funktioniert. Faktisch hast du jetzt eine [inaudible] oder Solar, die als Suchmaschinentechologie eingesetzt wird überall, und die werden genutzt. Aber: Jetzt gibt es dort Module, die professionelle Anwender, oder Firmen, schreiben oder anbieten lizenziert. Das ist das was ich vorhin meinte: Kümmert man sich jetzt wieder um den Use Case, auf den Use Case, und nicht die Kerntechnologie, Kerntechnologie ist für alle frei. Und auch das kann ich mir wieder vorstellen. Aber auch hier: Reine Open-Source Lösung glaube ich, das hat keine Zukunft. Also du brauchst immer irgendeinen Treiber, der dahinter steckt. Und zwar einen Treiber, der nachhaltig arbeiten kann, über Jahrzehnte das bewiesen hat. Also Daimler Chrysler würde kein – weiß ich nicht – MySQL einsetzen. Ja?

R [35:35-35:36]: Unwahrscheinlich.

E10 [35:36-35:37]: Genau.

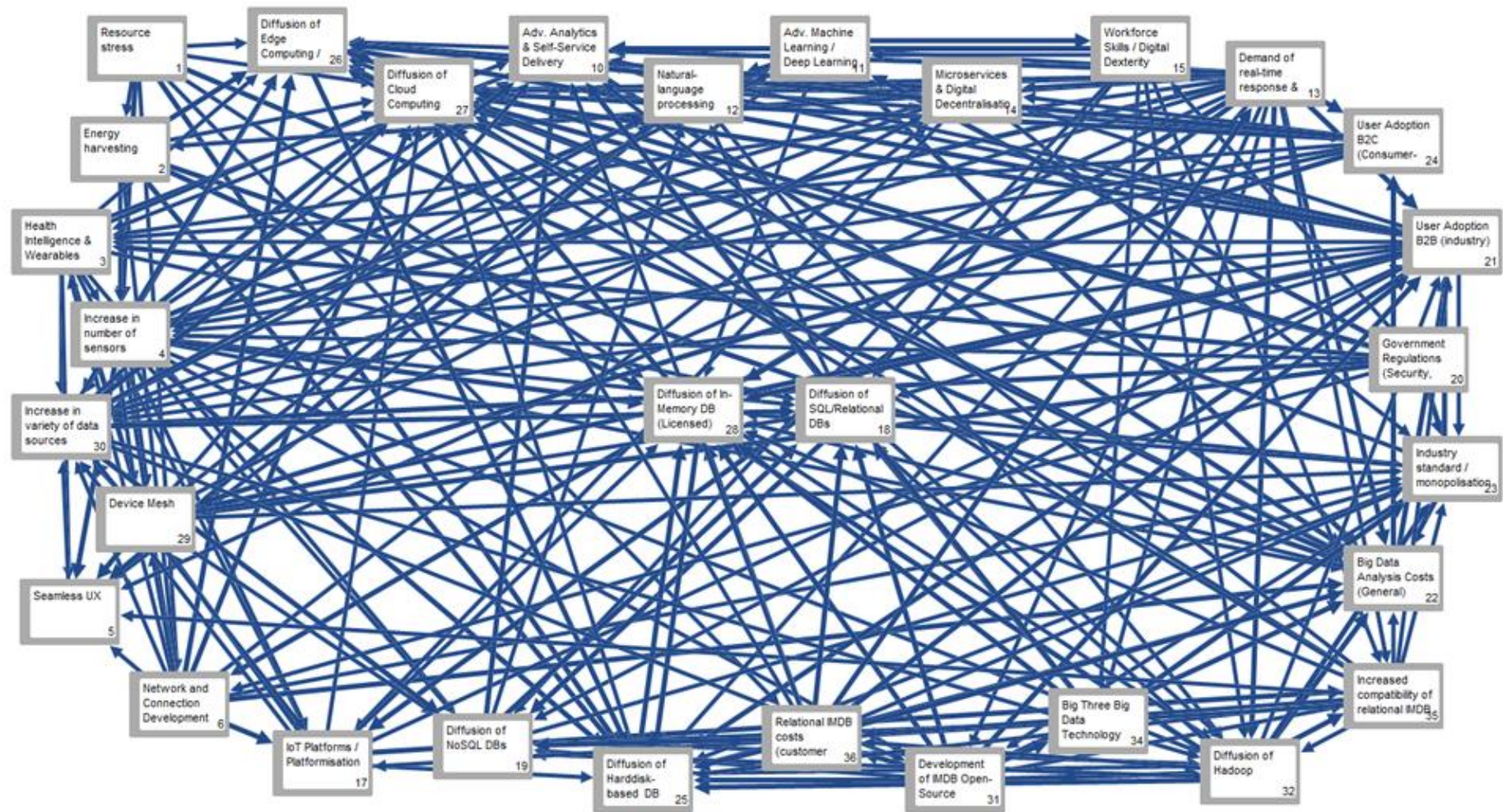
R [35:37-35:44]: OK. Cool. Gut, dann Dankeschön!

E10 [35:44-35:51]: Sehr sehr gerne. Ich hoffe ich konnte ein bisschen...

R [35:52-35:55]: Auf jeden Fall!

Attachment E: Parmenides EIDOS™ output files

Grid of 31 Drivers



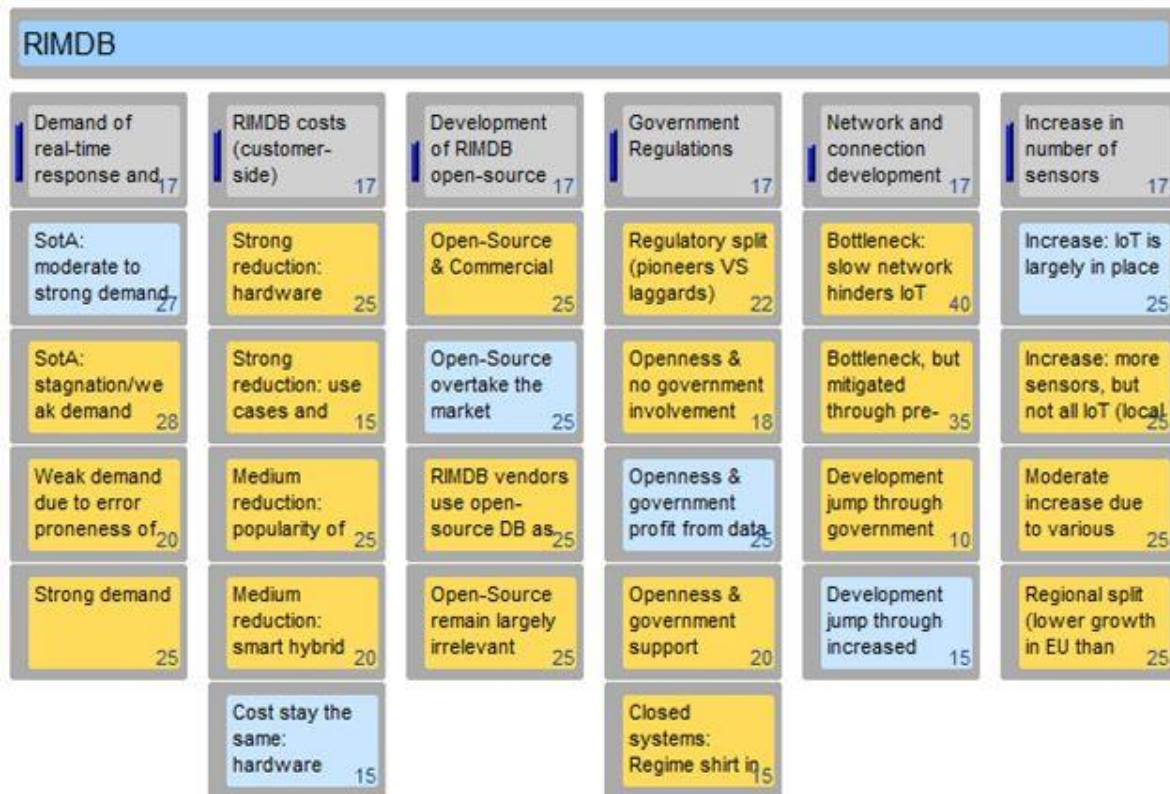
Active-Passive Map of Drivers

Highlighted: Key Drivers

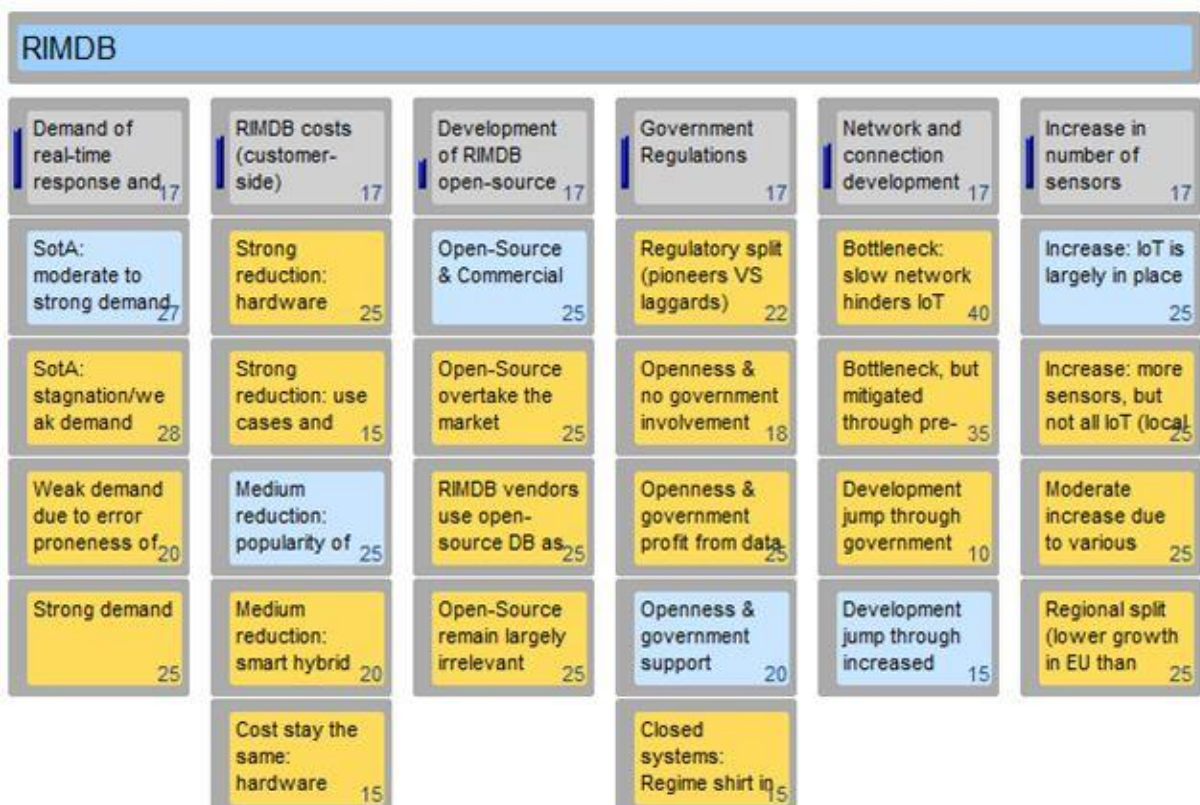


Initial Scenario Choice:

Scenario A



Scenario B



Scenario C

RIMDB					
Demand of real-time response and, 17	RIMDB costs (customer-side) 17	Development of RIMDB open-source 17	Government Regulations 17	Network and connection development 17	Increase in number of sensors 17
SotA: moderate to strong demand, 27	Strong reduction: hardware 25	Open-Source & Commercial 25	Regulatory split (pioneers VS laggards) 22	Bottleneck: slow network hinders IoT 40	Increase: IoT is largely in place 25
SotA: stagnation/weak demand 28	Strong reduction: use cases and 15	Open-Source overtake the market 25	Openness & no government involvement 18	Bottleneck, but mitigated through pre- 35	Increase: more sensors, but not all IoT (local) 25
Weak demand due to error proneness of 20	Medium reduction: popularity of 25	RIMDB vendors use open-source DB as 25	Openness & government profit from data 25	Development jump through government 10	Moderate increase due to various 25
Strong demand 25	Medium reduction: smart hybrid 20	Open-Source remain largely irrelevant 25	Openness & government support 20	Development jump through increased 15	Regional split (lower growth in EU than 25
	Cost stay the same: hardware 15		Closed systems: Regime shift in 15		

Scenario D

RIMDB					
Demand of real-time response and, 17	RIMDB costs (customer-side) 17	Development of RIMDB open-source 17	Government Regulations 17	Network and connection development 17	Increase in number of sensors 17
SotA: moderate to strong demand, 27	Strong reduction: hardware 25	Open-Source & Commercial 25	Regulatory split (pioneers VS laggards) 22	Bottleneck: slow network hinders IoT 40	Increase: IoT is largely in place 25
SotA: stagnation/weak demand 28	Strong reduction: use cases and 15	Open-Source overtake the market 25	Openness & no government involvement 18	Bottleneck, but mitigated through pre- 35	Increase: more sensors, but not all IoT (local) 25
Weak demand due to error proneness of 20	Medium reduction: popularity of 25	RIMDB vendors use open-source DB as 25	Openness & government profit from data 25	Development jump through government 10	Moderate increase due to various 25
Strong demand 25	Medium reduction: smart hybrid 20	Open-Source remain largely irrelevant 25	Openness & government support 20	Development jump through increased 15	Regional split (lower growth in EU than 25
	Cost stay the same: hardware 15		Closed systems: Regime shift in 15		

Final Scenario Choice: Rough Draft Projections

Scenario 1

Real-time response and service acceleration have become the state-of-the art. The demand for real-time analyses is moderate to strong.

Due to market pressures to adoption, even those firms who did not necessarily want to adopt real-time systems, or did not see the need for them, are forced to do so if they do not want to be kicked out of the market.

Larger corporations and those lagging behind in adoption of real-time response are still trying to catch up, but the development was not fast enough to kick all these laggards out of the market (yet).

Costs for relational in-memory database systems are not falling, but stagnated on the customer-side.

Although hardware-components continue to become cheaper overall, the exponentially larger data volumes require higher efficiency of the hardware. Thus, the increased efficiency of systems leads to stagnating prices.

Relational open-source database solutions take over the market.

One of the reasons for this development is that the large internet companies make a well-working relational in-memory database available to the community.

From government-side, a shift towards complete openness regarding the collection and use of data took place. Governments also want to benefit from this development.

Governments, especially those of the developed countries, have installed regulations which support the exploitation of big data for maximum benefits. However, these regulations also enable governments to take their piece of the Big-Data-pie and use available data for their own interests. One aspect of this development is industrial espionage.

The mobile network and overall connectivity has developed fast over the last years.

Several large-scale private investments boosted the network development. The effect of these investments was strengthened due to increased competition between incumbent network operators and the newcomer investors in the network and connectivity space.

The number of sensors has increased largely overall.

The actual development of the Internet of Things has played out very much in-line with projections of its growth. The large potentials in various fields are being reaped along the lines of "data is the new gold".

Scenario 2

Real-time response and service acceleration have become the state-of-the art. The demand for real-time analyses is moderate to strong.

Due to market pressures to adoption, even those firms who did not necessarily want to adopt real-time systems, or did not see the need for them, are forced to do so if they do not want to be kicked out of the market.

Larger corporations and those lagging behind in adoption of real-time response are still trying to catch up, but the development was not fast enough to kick all these laggards out of the market (yet).

Costs for relational in-memory database systems on customer-side have decreased significantly due to increasingly growing popularity of cloud-solutions.

Therefore, potentially large hardware investments fall away.

Both relational open-source as well as commercial relational in-memory database solutions exist next to each other.

Open-source solutions have caught up to commercial solutions regarding market-share, and set a high standard for all commercial products. Commercial products thus have to strive for adding extra value in order to stay competitive with open-source solutions.

From government-side, a shift towards complete openness regarding the collection and use of data took place. Governments support the exploitation of data for maximum economic value.

There are no regulations to handicap business around the collection and use of data. Governments even financially support various areas concerning big data to ensure optimal exploitation of value, e.g. by financing research and supporting the network and connection development

The mobile network and overall connectivity has developed fast over the last years.

Several large-scale private investments boosted the network development. The effect of these investments was strengthened due to increased competition between incumbent network operators and the newcomer investors in the network and connectivity space.

The number of sensors has increased largely overall.

The actual development of the Internet of Things has played out very much in-line with projections of its growth. The large potentials in various fields are being reaped along the lines of “data is the new gold”.

Scenario 3

Real-time response and service acceleration have become the state-of-the art. Yet, an overall demand in new systems is stagnating, i.e. weakened in comparison to previous years.

Those firms who did not implement technology for real-time analysis were kicked out of the market. The firms who are still in the market already have implemented technologies, and because improvements in response-speeds of systems are only marginal, demand is not as strong as it used to be. For the most part, firms newly starting up need to install real-time systems from scratch, and incumbents may need occasional updates.

Costs for relational in-memory database systems are not falling, but stagnated on the customer-side.

Although hardware-components continue to become cheaper overall, the exponentially larger data volumes require higher efficiency of the hardware. Thus, the increased efficiency of systems leads to stagnating prices.

The large vendors of previously commercial relational-in-memory databases now use these systems in an open-source form as a vehicle for selling other systems.

The vendor's focus shifted away from trying to make money with relational in-memory databases directly, but found more value in other technologies, for example relevant to the Internet of Things, or reorientation towards ERP- and CRM-systems.

App-store model

From government-side, a shift towards complete openness regarding the collection and use of data took place. Governments support the exploitation of data for maximum economic value.

There are no regulations to handicap business around the collection and use of data. Governments even financially support various areas concerning big data to ensure optimal exploitation of value, e.g. by financing research and supporting the network and connection development

The mobile network and overall connectivity has developed fast over the last years, because of a large-scale government investment effort.

The number of sensors has increased largely overall.

The actual development of the Internet of Things has played out very much in-line with projections of its growth. The large potentials in various fields are being reaped along the lines of “data is the new gold”.

References

- Accenture. (2016). *Accenture Technology Vision 2016: People First: The Primacy of People in a Digital Age*. Retrieved from https://www.accenture.com/t20160202T102002__w__/us-en/_acnmedia/Accenture/Omobono/TechnologyVision/pdf/Technology-Trends-Technology-Vision-2016.pdf
- Amazon Web Services, Inc. (n.d.). *Big Data Analytics: Erstellen, Sichern und nahtloses Skalieren von End-to-End-Anwendungen für Big Data in der Cloud*. Retrieved from <https://aws.amazon.com/de/big-data/>
- Ansoff, I. H. (1975). Managing Strategic Surprise by Response to Weak Signals. *California Management Review*, 18, 21–33.
- Atzori, L., & Iera, A. & Morabito, G. (2010). The Internet of Things: A survey. *Computer Networks*, 54(15), 2787–2805. doi:10.1016/j.comnet.2010.05.010
- Audi AG. (2015). *Audi Geschäftsbericht 2014: we create tomorrow*. Retrieved from <http://www.audi.de/gb2014>
- Barney, J. B. (1995). Looking inside for competitive advantage. *Academy of Management Perspectives*, 9(4), 49–61. doi:10.5465/AME.1995.9512032192
- Buhl, H. U., Röglinger, M., Moser, F. & Heidemann, J. (2013). Big Data. A Fashionable Topic with(out) Sustainable Relevance for Research and Practice? *Business & Information Systems Engineering*, 2, 65-69.
- Bundesministerium für Wirtschaft und Energie (2016). *Digitale Strategie 2025*. Retrieved from [http://www.bmwi.de/BMWi/Redaktion/PDF/Publikationen/digitale-strategie-2025,property=pdf,bereich=bmwi2012,sprache=de,rwb=true.pdf%29,](http://www.bmwi.de/BMWi/Redaktion/PDF/Publikationen/digitale-strategie-2025,property=pdf,bereich=bmwi2012,sprache=de,rwb=true.pdf%29)
- Burgelman, R. A., Christensen, C. M. & Wheelwright, S. C. (2009). *Strategic Management of Technology and Innovation*. New York: McGraw-Hill.
- Börjeson, L., Höjer, M., Dreborg, K.-H., & Ekvall, T. & Finnveden, G. (2006). Scenario types and techniques: Towards a user's guide. *Futures*, 38(7), 723–739. doi:10.1016/j.futures.2005.12.002
- Bradley, J., Loucks, J., Macaulay, J., Noronha, A., & Wade, M. (2015). *Digital Vortex: How Digital Disruption is Redefining Industries*. Retrieved from Global Center for Digital Business Transformation - an IMD and Cisco Initiative website: http://www.imd.org/uupload/IMD.WebSite/DBT/Digital_Vortex_06182015.pdf
- Brettel, M., Friederichsen, N., Keller, M., & Rosenberg, M. (2014). How Virtualization, Decentralization and Network Building Change the Manufacturing Landscape: An Industry 4.0 Perspective. *International Journal of Mechanical, Aerospace, Industrial, Mechatronic and Manufacturing Engineering*, 8(1).
- Bryman, A. & Bell, E. (2007). *Business Research Methods*: Oxford University Press. Retrieved from <https://books.google.de/books?id=vpqjq4I9KGM>
- Campbell-Hunt, C. (2000). What Have We Learned about Generic Competitive Strategy? A Meta-Analysis. *Strategic Management Journal*, 21(2), 127–154.

- Carter, J. (2015). *Check out 10 technologies that could change your life in 2016*. Techradar. Retrieved from <http://www.techradar.com/us/news/world-of-tech/future-tech/10-hugely-important-it-trends-for-2016-1308808>
- Carr, N. (2003). IT Doesn't Matter. *Harvard Business Review*. (May), 41–49. Retrieved from <http://web.a.ebscohost.com/ehost/pdfviewer/pdfviewer?sid=54783e32-ed2d-4276-be84-8e530d43b616%40sessionmgr4004&vid=0&hid=4206>
- Chapple, M. (n.d.). *The ACID Model*. Retrieved from <http://databases.about.com/od/specificproducts/a/acid.htm>
- Coates, J. F. (2010). The future of foresight—A US perspective. *Technological Forecasting and Social Change*, 77(9), 1428–1437. doi:10.1016/j.techfore.2010.07.009
- Codd, E. F. (1970). A relational model of data for large shared data banks. *Communications of the ACM*, 13(6), 377–387. doi:10.1145/362384.362685
- CSCMP & Sanders, N. (2014). *The definitive guide to manufacturing and service operations: Master the strategies and tactics for planning, organizing, and managing how products and services are produced*. Upper Saddle River, NJ: Pearson Education.
- Datawarehouse4u.info (n.d.). *OLTP vs. OLAP*. Retrieved from <http://datawarehouse4u.info/OLTP-vs-OLAP.html>
- Davis, J., Edgar, T., Porter, J., & Bernaden, J. Sarli, M. (2012). Smart manufacturing, manufacturing intelligence and demand-dynamic performance. *Computers & Chemical Engineering*, 47, 145–156. doi:10.1016/j.compchemeng.2012.06.037
- Daws, R. (2016). *Apple set to debut incredible LiFi technology*. TelecomsTech. Retrieved from <http://www.telecomstechnews.com/news/2016/jan/19/apple-set-debut-incredible-lifi-technology3/>
- Dean, J. & S. Ghemawat (2008) “MapReduce: Simplified Data Processing on Large Clusters”, *Communications of the Association for Information Systems*, (51)1, pp. 107–113
- Diaconu, C., Freedman, C., Ismert, E., Larson, P.-A., Mittal, P., Stonecipher, R., Verma, N. Mike, Z. (2013). *Hekaton: SQL Server's Memory - Optimized OLTP Engine*. Microsoft. Retrieved from <http://research.microsoft.com/pubs/193594/Hekaton%20-%20Sigmod2013%20final.pdf>
- Die Bundesbeauftragte für den Datenschutz und die Informationsfreiheit (n.d.). *Die Bundesbeauftragte für den Datenschutz und die Informationsfreiheit Home*. Retrieved from http://www.bfdi.bund.de/DE/Home/home_node.html
- Dolye, C. (2012). *Privacy: An Overview of the Electronic Communications Privacy Act*. Congressional Research Service. Retrieved from <https://www.fas.org/sgp/crs/misc/R41733.pdf>
- Drucker, P. F. (1986). *The Frontiers of Management: Where Tomorrow's Decisions Are Being Shaped Today* (1st ed.). New York: Truman Talley Books.
- Edlich, S. (n.d.). *NoSQL. Your Ultimate Guide to the Non-Relational Universe!* Retrieved from <http://nosql-database.org/>
- EIT Digital IVZW. (2016). *Innovation Radar: 2015 Annual Trend Report*. Brussels.

- European Commission (n.d.). *Protection of Personal Data*. Retrieved from <http://ec.europa.eu/justice/data-protection/>
- European Commission (2015a). *A Digital Single Market for Europe: Commission sets out 16 initiatives to make it happen*. Retrieved from http://europa.eu/rapid/press-release_IP-15-4919_en.htm
- European Commission (2015b). *Questions and Answers on the EU-US data protection "Umbrella agreement"*. Retrieved from http://europa.eu/rapid/press-release_MEMO-15-5612_en.htm
- European Commission (2016a). *EU Commission and United States agree on new framework for transatlantic data flows: EU-US Privacy Shield*. Retrieved from http://europa.eu/rapid/press-release_IP-16-216_en.htm
- European Commission (2016b). *Restoring trust in transatlantic data flows through strong safeguards: European Commission presents EU-U.S. Privacy Shield*. Retrieved from http://europa.eu/rapid/press-release_IP-16-433_en.htm?locale=en
- Fink, A., & Schlake, O. (2000). Scenario Management: An Approach for Strategic Foresight. *Competitive Intelligence Review*, 11(1), 37–45.
- Fjord. (2015). *FJORD TRENDS 2016: TRENDS IMPACTING DESIGN AND INNOVATION*. Retrieved from https://www.accenture.com/_acnmedia/PDF-1/Accenture-Fjord-Trends-2016.pdf
- Fleisch, E., Weinberger, M., & Wortmann, F. (2014). *Business Models and the Internet of Things: Bosch IoT Lab White Paper*. Retrieved from http://cocoa.ethz.ch/downloads/2014/10/2090_EN_Bosch%20Lab%20White%20Paper%20GM%20im%20IoT%201_2.pdf
- Lee, M. (2006). *Sun begins releasing Java under the GPL*. Free Software Foundation. Retrieved from <http://www.fsf.org/news/fsf-welcomes-gpl-java.html>
- Gartner, Inc. (2014a). *Gartner's 2014 Hype Cycle for Emerging Technologies Maps the Journey to Digital Business*. Retrieved from <http://www.gartner.com/newsroom/id/2819918>
- Gartner, Inc. (2014b). *Gartner Says 4.9 Billion Connected "Things" Will Be in Use in 2015: In 2020, 25 Billion Connected "Things" Will be In Use* Analysts to Explore the Disruptive Impact of IoT on Business at Gartner Symposium/ITxpo 2014, November 9-13 in Barcelona, Spain. Barcelona, Spain. Retrieved from <http://www.gartner.com/newsroom/id/2905717>
- Gartner, Inc. (2015a). *Gartner's 2015 Hype Cycle for Emerging Technologies Identifies the Computing Innovations That Organizations Should Monitor*. Retrieved from <http://www.gartner.com/newsroom/id/3114217>
- Gartner, Inc. (2015b). *Top 10 Technology Trends Signal the Digital Mesh: An evolving digital mesh of smart machines will connect billions of things into a continuous digital experience*. Retrieved from <http://www.gartner.com/smarterwithgartner/top-ten-technology-trends-signal-the-digital-mesh/>
- Gerpott, T. J. (2005). *Strategisches Technologie- und Innovationsmanagement*. 2nd Edition. Stuttgart: Schäffer-Poeschel.

- Gibbs, S. (2015). *What is 'safe harbour' and why did the EUCJ just declare it invalid?* The Guardian. Retrieved from <http://www.theguardian.com/technology/2015/oct/06/safe-harbour-european-court-declare-invalid-data-protection>
- Giusto, D., Atzori, L., Iera, A., & Morabito, G. (Eds.). (2010). *The Internet of Things: 20th Tyrrhenian Workshop on Digital Communications*. New York, NY: Springer Science+Business Media LLC.
- Godet, M., & Roubelat, F. (1996). Creating the future: the use and misuse of scenarios. *Long Range Planning*, 29(2), 164–171.
- Google, Inc. (2015a). *Third-party Tools and Services*. Retrieved from <https://cloud.google.com/bigquery/third-party-tools>
- Google, Inc. (2015b). *What is BigQuery?* Retrieved from <https://cloud.google.com/bigquery/what-is-bigquery>
- Grandpierre, M. & Buss, G. & Esser, R. (2013). *In-Memory Computing technology: The holy grail of analytics?* Retrieved from http://www2.deloitte.com/content/dam/Deloitte/de/Documents/technology-media-telecommunications/TMT_Studie_In_Memory_Computing.pdf
- Grant, R. M. (2010). *Contemporary Strategy Analysis: Text Edition* (7th ed.). Hoboken: Wiley.
- Grimes, C. A., Dickey E.C., & Pishko, M. V. (Eds.). (2005). *Encyclopedia of Sensors* (Vol. 9). Stevenson Ranch: American Scientific Publishers.
- Hallenbeck, C. (2015). *NoSQL, OldSQL, NewSQL, In-Memory & SAP HANA*. Retrieved from <https://blogs.saphana.com/2015/05/19/nosql-olddb-newsql-memory-sap-hana/>
- Heaton, B. (2015). *New York City Fights Fire with Data*. Analytics help New York City firefighters track potential hot spots. Emergency Management. Retrieved from <http://www.emergencymgmt.com/safety/New-York-City-Fights-Fire-Data.html>
- Helmrich, K. (2015). *Presentation: On the Way to Industrie 4.0 – The Digital Enterprise*. Retrieved from <http://www.siemens.com/press/pool/de/events/2015/digitalfactory/2015-04-hannovermesse/presentation-e.pdf>
- IBM Corporation (n.d.). *What is BLU Acceleration?* Retrieved from <http://www.ibmbluhub.com/why-blu-acceleration/>
- Jankowski, S. (2014). The Sectors Where the Internet of Things Really Matters. *Harvard Business Review*. (10), n/a. Retrieved from <https://hbr.org/2014/10/the-sectors-where-the-internet-of-things-really-matters/>
- Jankowski, S., Covello, J., Bellini, H., Ritchie, J. & Costa, D. (2014). *The Internet of Things: Making sense of the next mega-trend*. Retrieved from <http://www.goldmansachs.com/our-thinking/outlook/internet-of-things/iot-report.pdf>
- Javadi, B., Anderson, P. & Cappello, F. (2009). Cost-Benefit Analysis of Cloud Computing versus Desktop Grids. *Conference: Parallel & Distributed Processing, 2009*. doi:10.1109/IPDPS.2009.5160911

- Koch, V., Kuge, S., Geissenbauer, R., & Schrauf, S. (2014). *Industry 4.0: Opportunities and challenges of the industrial internet*. Retrieved from <http://www.strategyand.pwc.com/media/file/Industry-4-0.pdf>
- Kopetz, H. (2011). *Real-Time Systems*. Boston, MA: Springer US.
- KPMG. (2014). *Future State 2030: The global megatrends shaping governments*. Retrieved from <https://www.kpmg.com/Global/en/IssuesAndInsights/ArticlesPublications/future-state-government/Documents/future-state-2030-v3.pdf>
- Krueger, J., Grund, M., Tinnefeld, C., Eckart, B., Zeier, A. & Plattner, H. (2010). Hauptspeicherdatenbanken für Unternehmensanwendungen. *Datenbank-Spektrum*, 10(3), 143–158. doi:10.1007/s13222-010-0031-5
- LaValle, S., Lesser, E., Shockley, R., Hopkins, M. S. & Kruschwitz, N. (2011). Big Data, Analytics and the Path from Insights to Value. *MITSloan Management Review*, 52(2), 20-32.
- Lindström, J., Raatikka, V., Ruuth, J., Soini, P. & Vakkila, K. (2013). *IBM solidDB: In-Memory Database Optimized for Extreme Speed and Availability*. Retrieved from <http://sites.computer.org/debull/A13june/soliddb1.pdf>
- Little International, A. D. (Ed.) (1985). *Management im Zeitalter der strategischen Führung*. Wiesbaden: Gabler.
- Löffler, M. & Tschiesner, A. (2013). *The Internet of Things and the future of manufacturing*. Retrieved from <http://www.futurenautics.com/wp-content/uploads/2013/10/Internet-of-Things-and-future-of-manufacturing.pdf>
- Loos, P., Lechtenbörger, J., Vossen, G., Zeier, A., Krüger, J., Müller, J. & Winter, R. (2011). In-memory Databases in Business Information Systems. *Business & Information Systems Engineering*, 3(6), 389–395. doi:10.1007/s12599-011-0188-y
- Lycett, M. (2013). ‘Datafication’: making sense of (big) data in a complex world. *European Journal of Information Systems*, 22, 281-386. doi:10.1057/ejis.2013.10
- Manyika, J., Chui, M., Bisson, P., Woetzel, J., Dobbs, R., Bughin, J. & Aharon, D. (2015). *The Internet of Things: Mapping The Value Beyond the Hype*. Retrieved from http://www.mckinsey.com/~media/McKinsey/dotcom/Insights/Business%20Technology/Unlocking%20the%20potential%20of%20the%20Internet%20of%20Things/Unlocking_the_potential_of_the_Internet_of_Things_Full_report.ashx
- Marston, S., Li, Z., Bandyopadhyay, S., & Zhang, J. & Ghalsasi, A. (2011). Cloud computing — The business perspective. *Decision Support Systems*, 51(1), 176–189. doi:10.1016/j.dss.2010.12.006
- Martin, B. R. (2010). The origins of the concept of ‘foresight’ in science and technology: An insider's perspective. *Technological Forecasting and Social Change*, 77(9), 1438–1447. doi:10.1016/j.techfore.2010.06.009
- Mastrianna, F. V. (2013). *Basic Economics: 16th Edition*. Mason: South-Western.
- Mazhelis, O., & Luoma, E. & Warma, H. (Eds.). (2012). *Internet of Things, Smart Spaces and Next Generation Networking. Lecture Notes in Computer Science: Vol. 7469. Defining an Internet-of-Things Ecosystem*. Heidelberg: Springer-Verlag.

- McAfee, A. & Brynjolfsson, E. (2012). Big Data: The Management Revolution. *Harvard Business Review*, 10, 59-68.
- Mietzner, D., & Reger, G. (2005). Advantages and disadvantages of scenario approaches for strategic foresight. *International Journal of Technology Intelligence and Planning*, 1(2), 220-239. doi:10.1504/IJTIP.2005.006516
- Mietzner, D. (2009). *Strategische Vorausschau und Szenarioanalysen*. Wiesbaden: Gabler.
- Miles, I. (2010). The development of technology foresight: A review. *Technological Forecasting and Social Change*, 77(9), 1448-1456. doi:10.1016/j.techfore.2010.07.016
- Mintzberg, H. (1978). Patterns in Strategy Formation. *Management Science*, 24(9), 934-948.
- Mintzberg, H. (1987). Crafting Strategy. *Harvard Business Review*, 7, 66-75.
- Mintzberg, H. (1994). The Fall and Rise of Strategic Planning. *Harvard Business Review*, 1, 109-114.
- Miorandi, D., Sicari, S., & de Pellegrini, F. & Chlamtac, I. (2012). Internet of things: Vision, applications and research challenges. *Ad Hoc Networks*, 10(7), 1497-1516. doi:10.1016/j.adhoc.2012.02.016
- OECD (Ed.) (2006). *Schooling for tomorrow. Think scenarios, rethink education*. Paris: OECD.
- Oracle Corporation (n.d.). *TimesTen In-Memory Database and TimesTen Application-Tier Database Cache*. Retrieved from <http://www.oracle.com/technetwork/database/database-technologies/timesten/overview/index.html>
- Oracle Corporation (2015). *Oracle Database In-Memory: Oracle White Paper*. Retrieved from <http://www.oracle.com/technetwork/database/in-memory/overview/twp-oracle-database-in-memory-2245633.html>
- Plattner, H. & Leukert, B. (2015). *The In-Memory revolution: How SAP HANA enables business of the future*. Cham: Springer.
- Porter, M. E. (1985a). *Competitive Advantage: Creating and Sustaining Superior Performance*. New York: The Free Press.
- Porter, M. E. (1985b). TECHNOLOGY AND COMPETITIVE ADVANTAGE. *Journal of Business Strategy*, 5(3), 60-78. doi:10.1108/eb039075
- Porter, M. E. (1980). *Competitive Strategy: Techniques for Analyzing Industries and Competitors*. New York: The Free Press.
- Porter, M. E. (1996). What is Strategy? *Harvard Business Review*, 11, 59-78.
- Porter, M. E. (1997). COMPETITIVE STRATEGY. *Measuring Business Excellence*, 1(2), 12-17. doi:10.1108/eb025476
- Porter, M. E. (2008). *On Competition: Updated and Expanded Edition*. A Harvard Business Review Book. Boston, MA: Harvard Business School Publishing.
- Porter, M. E. (2008). The Five Competitive Forces that Shape Strategy. *Harvard Business Review*, 86(01), 79-93. Retrieved from

<http://web.a.ebscohost.com/ehost/pdfviewer/pdfviewer?sid=76fc15b4-0d29-4576-a2fb-c1ffd8743125%40sessionmgr4005&vid=4&hid=4201>

- Porter, M. E. & Heppelmann, J. E. (2014). How Smart, Connected Products are Transforming Competition. *Harvard Business Review*. (11), n/a. Retrieved from <https://hbr.org/2014/11/how-smart-connected-products-are-transforming-competition>
- Porter, M. E. & Millar, V. E. (1985). How Information Gives You Competitive Advantage. *Harvard Business Review*, 63(4), 149–160.
- Poynder, R. (2001). The Open Source Movement Does this software provide a viable, user-friendly alternative to proprietary solutions? *InformationToday*, 18(9). Retrieved from <http://www.infoday.com/it/oct01/poynder.htm>
- Press, G. (2014). *It's Official: The Internet Of Things Takes Over Big Data As The Most Hyped Technology*. Retrieved from <http://www.forbes.com/sites/gilpress/2014/08/18/its-official-the-internet-of-things-takes-over-big-data-as-the-most-hyped-technology/>
- pureLiFi (2014). *What is Li-Fi?* Retrieved from http://purelifi.com/what_is_li-fi/
- Reibnitz, U. von. (1988). *Scenario Techniques*. New York: McGraw-Hill Publishing Co.
- Robbins, S. P., & Coulter, M. K. (2012). *Management: 11th Edition* (11th ed.). Boston: Prentice Hall.
- Rohrbeck, R. (2012). Exploring value creation from corporate-foresight activities. *Futures*, 44(5), 440–452. doi:10.1016/j.futures.2012.03.006
- Rohrbeck, R., Arnold, H. M., & Heuer, J. (2007). *Strategic Foresight in multinational companies a case study on the Deutsche Telekom Laboratories*. ISPIM-Asia 2007 conference, New Delhi, India.
- Rohrbeck, R., & Bade, M. (2012). *Environmental scanning, futures research, strategic foresight and organizational future orientation: a review, integration, and future research directions*. ISPIM Annual Conference 2012, Barcelona, Spain.
- Rohrbeck, R., Battistella, C., & Huizingh, E. (2015). Corporate foresight: An emerging field with a rich tradition. *Technological Forecasting and Social Change*, 101, 1–9. doi:10.1016/j.techfore.2015.11.002
- Rohrbeck, R., & Gemünden, H. G. (2008). *Strategic Foresight in Multinational Enterprises: Building a Best-Practice Framework from Case Studies*. R&D Management Conference 2008 "Emerging methods in R&D management", Ottawa, Canada.
- Rohrbeck, R., & Gemünden, H. G. (2011). Corporate foresight: Its three roles in enhancing the innovation capacity of a firm. *Technological Forecasting and Social Change*, 78(2), 231–243. doi:10.1016/j.techfore.2010.06.019
- Rohrbeck, R., & Schwarz, J. O. (2013). The value contribution of strategic foresight: Insights from an empirical study of large European companies. *Technological Forecasting and Social Change*, 80(8), 1593–1606. doi:10.1016/j.techfore.2013.01.004
- SAP America, Inc. (n.d., a). *SAP HANA Home*. Retrieved from <https://hana.sap.com/abouthana.html>
- SAP America, Inc. (n.d., b). *Gain unprecedented insight with Hadoop and the SAP HANA platform for Big Data*. Retrieved from <http://www.sap.com/solution/big-data/software/hadoop/index.html>

- Sato, K. (2012). *An Inside Look at Google BigQuery: White Paper / BigQuery*. Retrieved from <https://cloud.google.com/files/BigQueryTechnicalWP.pdf>
- Saunders, M., Lewis, P., & Thornhill, A. (2009). *Research methods for business students* (5. ed.). Harlow: Financial Times Prentice Hall.
- Schoemaker, P. (1995). Scenario planning: A tool for strategic thinking. *Sloan Management Review*, 36(2), 25–40. doi:10.1016/0024-6301(95)91604-0
- Schuh, G., & Klappert, S. (Eds.). (2011). *Technologiemanagement: Handbuch Produktion und Management 2* (2nd ed.). Heidelberg: Springer.
- Sommerlatte, T. & Deschamps, J.-P. (1985). Der strategische Einsatz von Technologien Konzepte und Methoden zur Einbeziehung von Technologien in die Strategieentwicklung des Unternehmens. In Arthur D. Little International (Ed.). *Management im Zeitalter der Strategischen Führung* (pp.39-74). Wiesbaden: Gabler.
- Spanos, Y. E., & Lioukas, S. (2001). An examination into the causal logic of rent generation: Contrasting Porter's competitive strategy framework and the resource-based perspective. *Strategic Management Journal*, 22(10), 907–934. doi:10.1002/smj.174
- Specht, D. & Möhrle, M. G. (Eds.) (2002). *Gabler Lexikon Technologie Management*. Wiesbaden: Gabler Verlag.
- Stabell, C. B., & Fjeldstad, O. D. (1998). Configuring Value for Competitive Advantage: On Chains, Shops and Networks. *Strategic Management Journal*, 19(5), 413–437.
- Strebel, H. (Ed.). (2007). *UTB: Vol. 2455. Innovations- und Technologiemanagement*. Wien: Facultas-WUV.
- Swan, M. (2012). Sensor Mania! The Internet of Things, Wearable Computing, Objective Metrics, and the Quantified Self 2.0. *Journal of Sensor and Actuator Networks*, 1(3), 217–253. doi:10.3390/jsan1030217
- Techopedia (2016). Relational Database Management System (RDBMS). Retrieved from <https://www.techopedia.com/definition/1235/relational-database-management-system-rdbms>
- TED Conferences, LLC (n.d.). *Alan Kay. Educator and computing pioneer*. Retrieved from https://www.ted.com/speakers/alan_kay
- The Apache Software Foundation (n.d.). *Spark™. Lightning-fast cluster computing*. Retrieved from <http://spark.apache.org/>
- ThoughtWorks (2015). *TECHNOLOGY RADAR NOV '15: Our thoughts on the technology and trends that are shaping the future*.
- van Notten, P. W. F., Rotmans, J., van Asselt, M. B. A., & Rothman, D. S. (2003). An updated scenario typology. *Futures*, 35(5), 423–443. doi:10.1016/S0016-3287(02)00090-3
- van Notten, P. W. F. (2006). Chapter 4: Scenario development: a typology of approaches. In OECD (Ed.), *Schooling for tomorrow. Think scenarios, rethink education*. Paris: OECD. Retrieved from <http://www.oecd.org/site/schoolingfortomorrowknowledgebase/futuresthinking/scenarios/37246431.pdf>

- van Notten, P. W. F., Slegers, A. M., & van Asselt, M. B. A. (2005). The future shocks: On discontinuity and scenario development. *Technological Forecasting and Social Change*, 72(2), 175–194. doi:10.1016/j.techfore.2003.12.003
- Vecchiato, R. (2015). Creating value through foresight: First mover advantages and strategic agility. *Technological Forecasting and Social Change*, 101, 25–36. doi:10.1016/j.techfore.2014.08.016
- Vom Brocke, J., Debortoli, S., & Müller, O. & Reuter, N. (2014). How In-memory Technology Can Create Business Value: Insights from the Hilti Case. *Communications of the Association for Information Systems*, 34(7), 151–168.
- Walsh, S. T., & Kirchhoff, B. A. & Newbert, S. (2002). Differentiating market strategies for disruptive technologies. *IEEE Transactions on Engineering Management*, 49(4), 341–351. doi:10.1109/TEM.2002.806718
- Wernerfelt, B. (1984). A Resource-Based View of the Firm. *Strategic Management Journal*, 5(2), 171–180.
- Wernerfelt, B., & Karnani, A. (1987). COMPETITIVE STRATEGY UNDER UNCERTAINTY. *Strategic Management Journal*, 8(2), 187–194.
- Zahra, S. A. (1996). Technology Strategy and Financial Performance: Examining the Moderating Role of the Firm's Competitive Environment. *Journal of Business Venturing*, 11, 189–219.
- Zook, J. D., & Schroeder, N. (2005). Sensors as Information Transducers. In C. A. Grimes, Dickey E.C., & M. V. Pishko (Eds.), *Encyclopedia of Sensors* (Vol. 9, pp. 329–359). Stevenson Ranch: American Scientific Publishers.