DISKUSSIONSBEITRÄGE DER
FAKULTÄT FÜR BETRIEBSWIRTSCHAFTSLEHRE
MERCATOR SCHOOL OF MANAGEMENT
UNIVERSITÄT DUISBURG-ESSEN

Nr. 404

Cross-Company Business Intelligence: Definition, Model, and Quality Measurement

Jens Kaufmann

Universität Duisburg-Essen
Fakultät für Betriebswirtschaftslehre
Mercator School of Management
Lehrstuhl für Wirtschaftsinformatik, insbesondere Business Intelligence
Lotharstraße 63
D-47057 Duisburg
Telefon: (+49)203 / 379 - 2627
Telefax: (+49)203 / 379 - 1856
E-Mail: sekretariat.wi@uni-duisburg-essen.de
URL: www.wi.msm.uni-due.de

August 2016

Abstract

Business Intelligence is a well-established term for methods, concepts and tools to retrieve, store, deliver and analyze data for management and business purposes. Although collaboration across company borders has substantially increased over the past decades, little research has been conducted specifically on Cross-Company BI. Here we propose a working definition and distinction from general collaborative decision making. We create a model that takes existing research and related approaches of adjacent fields into account as well as a peer-to-peer network design. With an extensive simulation and parameter testing we show that the design proves valuable and competitive to centralized approaches and that obtaining a critical mass of participants leads to improved usefulness of the network. To quantify the observations, we introduce appropriate quality measures rigorously derived from respected concepts on data and information quality and multidimensional data models.

Keywords

Business Intelligence
Collaborative Business Intelligence
Cross-Company Business Intelligence
Quality Measurement

Index

List of figures

List of tables

1 Introduction

Today's business is highly interconnected and operating on a high technical niveau. Decreasing costs and latency times of communication, information gathering and analysis lead to a growing need and possibility of decentralized decision authority in companies (Bloom et al., 2014). Thus, exchange of data and information is constantly rising and external data sources can be integrated in the process. The gathering, provisioning and analyzing of data is most often subsumed under the term Business Intelligence (BI), which covers concepts, methods, and systems of decision support in enterprises (Watson and Wixom, 2007). While there has been done extensive research on inter-company processes, e.g., around the well-known concept of Supply Chain Management (Mentzer et al., 2001), cooperative work across company borders in BI has sparsely been considered and therefore is a desired field of research in Information Systems (IS) (Baars et al., 2014). Applications of these cross-company relations can, e.g., be found in the sectors of research, industry, and health, and are especially valuable for small and medium-sized enterprises (cf. section 2.3). Existing approaches are often based on the premise of a central 'data scheme' or even a central system (e.g., Martins et al. (2012); Mettler and Raber (2011)). The possible advantages of fully decentralized systems, first and foremost independence of the participants and flexible structures (Schoder and Fischbach, 2003), so far have not been taken seriously into account for research on peer-to-peer(P2P)-based BI networks, although P2P networks have proven useful for general data and information exchange (Barkai, 2001; Miller, 2001; Schoder and Fischbach, 2002). A first working model of P2P-based BI has been given by Golfarelli et al. (2011) with the Business Intelligence Network (BIN). The authors focus on technical terms, especially on reformulations of database queries in those networks, and give a broadened research agenda which describes the need to develop a more comprehensive model, e.g., for a discussion of security topics, and to discover quality measures that can be used to determine the usefulness of the concept. The authors use the term Collaborative BI (CBI) (Golfarelli et al., 2012a), to describe their field of work. This term has been identified as being not (solely) focused on inter-company work (Kaufmann and Chamoni, 2014).

We therefore aim to achieve and provide an understanding of (1) Cross-Company BI (CCBI) as a more suitable term and concept, (2) the creation and structure of CCBI-networks, (3) how their quality or usefulness can be measured, and (4) in which configurations they can be used for practical needs in an inter-company context. To do so, we choose a design science research approach, which focuses on artefact creation (in our case: a model and a prototypical implementation of a network simulation tool) to gain insight (Hevner et al., 2004). More precisely, we work on an exaptation (Gill and Hevner, 2013), i.e., we extend known ideas and methods (network design, BI systems) to a new domain (cross-company BI). By that we deliver prescriptive knowledge on CCBI-networks and the respective simulation models as well as descriptive knowledge, e.g., by evaluating the idea of a 'critical mass', needed for successful network usage (Katz and Shapiro, 1994), to these networks.

The remainder of this article is structured as follows. Chapter 2 covers the state of the art on CBI and CCBI to give a working definition of the topic. Chapter 3 describes relevant aspects for a comprehensive model of CCBI-networks and the model itself. Chapter 4 uses the model to develop applicable quality measures. By using an implementation of a simulation tool for CCBI networks, Chapter 5 demonstrates the appropriability of the quality measures themselves as well as the results of different studies on varying network configurations which determine guidelines for the creation of practically usable networks. Our conclusions and a view on future research are given in Chapter 6.[1]

---

[1] This article is based on a dissertation thesis and partly only shows an excerpt of more comprehensive work (cf. Kaufmann (2015)).

2 Towards a definition of Cross-Company Business Intelligence

This chapter gives a brief introduction into the field of BI in general and then presents the methodology and results of a literature review on collaborative and cross-company BI. The findings are used to create a working definition of cross-company BI and to distinguish different understandings of the various terms and applications.

2.1 Business Intelligence and Data Warehousing

BI systems are understood as data-driven Decision Support Systems (DSS), as all parts are based on data gathering, provision, and analysis or presentation (Laudon and Laudon, 2006; Power, 2009; Sharda et al., 2014). The main concept or tool in BI is the Data Warehouse (DW), which we understand in the form described by Inmon (2005): “A data warehouse is a subject-oriented, integrated, nonvolatile, and time-variant collection of data in support of management’s decisions.“ (Inmon, 2005, p. 29) The stored data is usually multi-dimensional and can be represented by a (hyper-)cube and analyzed with the means of Online Analytic Processing (OLAP). Every cube is spanned by a number of dimensions, consisting of elements that may be hierarchically organized. The combination of elements along all given dimensions then leads to the point of stored data, the ‘fact’. We use the $\mathcal{MD}$-notation as an accepted standard (Vaisman et al., 2009) to provide a basic formal definition of this concept for the purpose of measure definitions later on. We reference to Cabibbo and Torlone (1998) and Torlone (2008) for a more detailed version of our short explanation:

A *dimension scheme* $S(d)$ consists of a set of hierarchy levels $L = \{l_1, \dots, l_n\}$ and a partial order $\preccurlyeq$ on L, with $l_1 \preccurlyeq l_2$ meaning: Elements in l_1 are aggregated to elements in l_2, saying they are ‘rolled up’. L is a finite set and has a bottom element $\perp$ (regarding $\preccurlyeq$).

A *dimension instance* $I(d)$ consists of a function m, that associates real world elements to levels and a group of functions ρ, having a roll-up function $\rho^{l_1 \to l_2}: m(l_1) \to m(l_2)$ for all pairs $l_1 \preccurlyeq l_2$.

A *dimension* d is build up by a scheme $S(d)$ and an instance $I(d)$.

A *cube* c consists of a set of dimensions $D = \{d_1, \dots d_n\}$, and a (at this point not detailed) fact scheme and instance function, assigning fact values to coordinates, i.e., tuples of elements.

We will use the exemplary dimension ‘geography’ (see Fig. 1), consisting of three levels with one, three or six elements respectively, throughout the paper to describe our considerations on elements, levels, and their use for quality measure definitions.

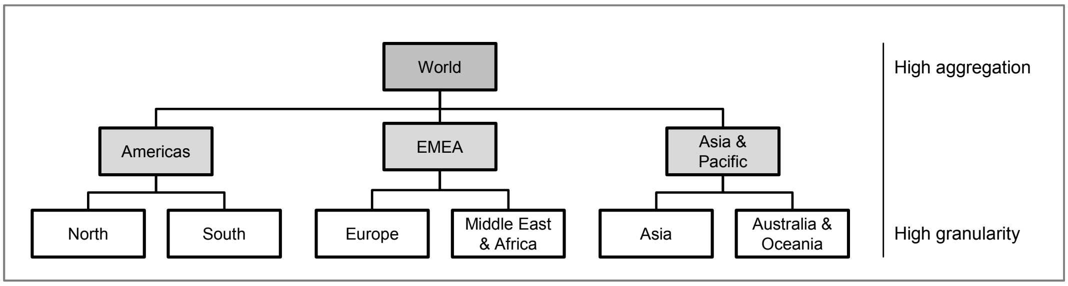


Fig. 1: Exemplary 'geography' dimension with hierarchy

Besides creating ad-hoc queries on the data warehouse and OLAP cubes, the stored data is used as a data base for reporting and various tasks in the context of Knowledge Discovery in Databases (KDD), especially making use of data mining algorithms to identify and apply useful patterns in the data by using techniques such as classification, clustering, and association rule mining (Fayyad et al., 1996). BI systems can be considered to be an industry standard as of today, with their analytical applications still gaining momentum due to ongoing developments in the fields of Business Analytics and Big Data (Forbes, 2015; Gartner Inc., 2016).

2.2 Collaborative BI

There has been extensive research about the support of decision making processes in groups by DSS and information technology (IT) in general (DeSanctis and Gallupe, 1987; Markus and Silver, 2008). The respective systems are, in difference to data-driven BI systems, classified as communication-driven DSS (Power, 2009) and support the decision making process itself rather than providing collaborative features in a BI context. This is changed by the prominent availability and use of Social Media, i.e., tools for simple messaging and networking via the Internet. Over the past few years, BI systems have incorporated these tools to provide easy accessible solutions for communication of different analysts as described in one of the first substantial publications on the topic by Dayal et al. (2008). This 'collaborative BI' thereafter has not only been discussed in research, but also in market studies, e.g., by Gartner Inc. (2011, 2014) in the context of Collaborative Decision Making. In a former literature review on CBI (as of March 2013) we have shown that the term is not used unambiguously (Kaufmann & Chamoni, 2014). Instead, three different understandings could be identified that focus on different combinations of internal and external data sources and data analysts. As this review provides a solid base for our considerations on company-internal and company-spanning (or 'cross-company') collaboration in the field of BI, we created an updated literature review on the field of Collaborative BI. We followed the method as given by the previous review to renew the findings and to determine the state of the art. We thereby covered all publications of the AIS senior scholars' basket of eight (AIS, 2011), the design science research oriented journal and conference list provided by the German WKWI (2008), and all relevant publications in an appropriate backward search (Webster and Watson, 2002). Tab. 1 compares the results of both of the reviews by showing the number of publications for each category using the original distinction into three fields of CBI. Additionally, following Böhringer et al. (2010) and Turban et al. (2011), for each category it is shown whether the authors focus on technical, organizational, functional, and/or economic aspects. This illustrates the strong focus on technical aspects that most authors use to define CBI.

Tab. 1: Results of the literature review on CBI

Category	# publications	Aspects discussed			
	2015 (2013)	Technical	Organizational	Functional	Economic
Internal Communication	36 (20)	33 (17)	7 (6)	2 (2)	6 (4)
Partnership in data	8 (4)	7 (2)	1 (2)	5 (2)	6 (2)
Partnership in analysis	4 (3)	3 (2)	1 (1)	0 (0)	4 (3)

While all identified references use the term Collaborative BI, it can clearly be seen that the majority of authors uses it in an understanding that focuses on company-internal communication which in some cases can include communication with partner organizations, e.g., for getting additional advice, but never includes data sharing. These authors mainly use CBI to describe the use of Social Media-like functions to provide analysts with tools to collaborate in existing BI environments. Authors that understand the term in a cross-company meaning either focus on partnership in data, where data sources are shared, but then used by every company on its own, or partnership in analysis, where data is shared and then is analyzed in a combined effort, e.g., to profit from different data models in the insurance industry.

2.3 Definition and review of Cross-Company BI

Considering the results of the current literature review on Collaborative BI, we argue that a modern definition should reflect the strong focus on internal collaboration. Such a definition is given by Muntean (2012), who is one of the authors in the reference list: “Collaborative Business Intelligence – the integration of information sharing features and functionality of popular Web 2.0 technologies and social media platforms within a BI platform (…)“ (Muntean, 2012, p. 196). We therefore also argue that a different term should be used for all the work on collaboration in BI that extends over company borders. Besides the usage of ‘Collaborative BI’ ((Alwis, 2004; Golfarelli et al., 2010, 2011, 2012a; Klarmann et al., 2013; Liu and Daniels, 2012; Martins et al., 2012; Mettler & Raber, 2011; Rizzi, 2012; Schwalm and Bange, 2004; Vera-Baquero et al., 2013; Werner et al., 2010)), our review has brought up the phrases “BI across enterprise borders” (Baars et al., 2014, p. 13), “information sharing across organizational borders” (Gartner Inc., 2013) , and “cross-enterprise business intelligence” (Simon and Shaffer, 2002, p. 132). We conducted a second literature review using Google Scholar to identify any publication that would make use of one of these terms and also added ‘company-spanning’, ‘inter(-)company’, ‘inter(-)organizational’ and ‘cross-company’ as possible search terms in combination with ‘Business Intelligence’ to see if any accepted definition of the concept had not been identified by the literature review before. The search terms were comprised of the phrases for denoting a company-spanning intention plus ‘Business Intelligence’ and making use of the AROUND(20)-operator to find every combination of these phrases within a 20-word range. Tab. 2 shows the number of results for each search term. The references have been added if they (1) clearly addressed the topic, (2) were not found in the former literature search, (3) were not coextensive publications by the same authors previously identified.

Tab. 2: Results of the literature review on cross-company terms and BI

Search term	# results	References found
across company borders	1	(Wuertz et al., 2013)
across enterprise borders	0	
across organizational borders	0	
cross enterprise	13	(Arigliano et al., 2013), note that this publication accounts for 9 of the 13 results
cross company	0	
company spanning	0	
enterprise spanning	0	
inter-company	0	
intercompany	1	(Rodriguez et al., 2010)
inter-organizational	6	
interorganizational	10	

Wuertz et al. (2013) argue that for a conglomerate of companies working together, a BI organization, e.g., a company-spanning BICC must be established. They give the example of a product-oriented DWH for industrial companies to share information, but do not focus on the general concept of cross-company BI. Arigliano et al. (2013) use the example of Italian industrial SMEs to discuss the need of integration in decision making and BI without giving further details. Rodriguez et al. (2010) refer to intercompany data distribution in the healthcare sector, but do not give any definition of that term or discuss it further. All three articles can be understood as focusing more on partnership in data than in analysis. While arguably Google Scholar is not the one-and-only reference for a literature review, it has been shown that its flaws should be considered to be on the side of precision rather than on the coverage of articles, which in fact proves to be very high (Bramer et al., 2016; Harzing, 2012; Meier and Conkling, 2008). Therefore we argue that with a high certainty there is no widely accepted term or definition for the concept in question and give

Definition (1) ***Cross-Company Business Intelligence (CCBI)*** describes concepts, methods, and systems to support data gathering, storage, provision, presentation, and analysis across company borders with regards to the independence of all participants.

We emphasize the latter part of the definition to focus on the willing collaboration of companies and to prevent confusion with BI systems that are just provided via 'the cloud' (Baars and Kemper, 2010), centralized systems where one company holds all data of all partners (Mettler & Raber, 2011), or 'Information as a service'-approaches where data and information access is sold by a single instance (Gartner Inc., 2013). As a practical use case we refer to the idea of loosely coupled institutions that share data as universities might when evaluating research activities (Golfarelli et al., 2012a). External links and collaborative work, for example, but not limited to, in decision support, have also been of special interest for small and medium enterprises in research and development (R&D) activities (Ale Ebrahim et al., 2009; Rothwell and Dodgson, 1991). A practical implementation of those company networks that emphasizes the independence of the partners and allows for project-bound connecting and disconnecting can be found in the German automotive sector (Scholta, 2005), although no ready-to-use CCBI-network has been implemented there yet. Finally, Fig. 2 shows the orientation framework used in (Kaufmann & Chamoni, 2014) to clearly distinguish CBI and CCBI. Identified publications of both literature reviews are shown as filled squares and, where applicable, are grouped by rectangles to show a similar understanding. Publications that only deal with related concepts are not shown in this overview, but are identified and analyzed in section 3 of this article.

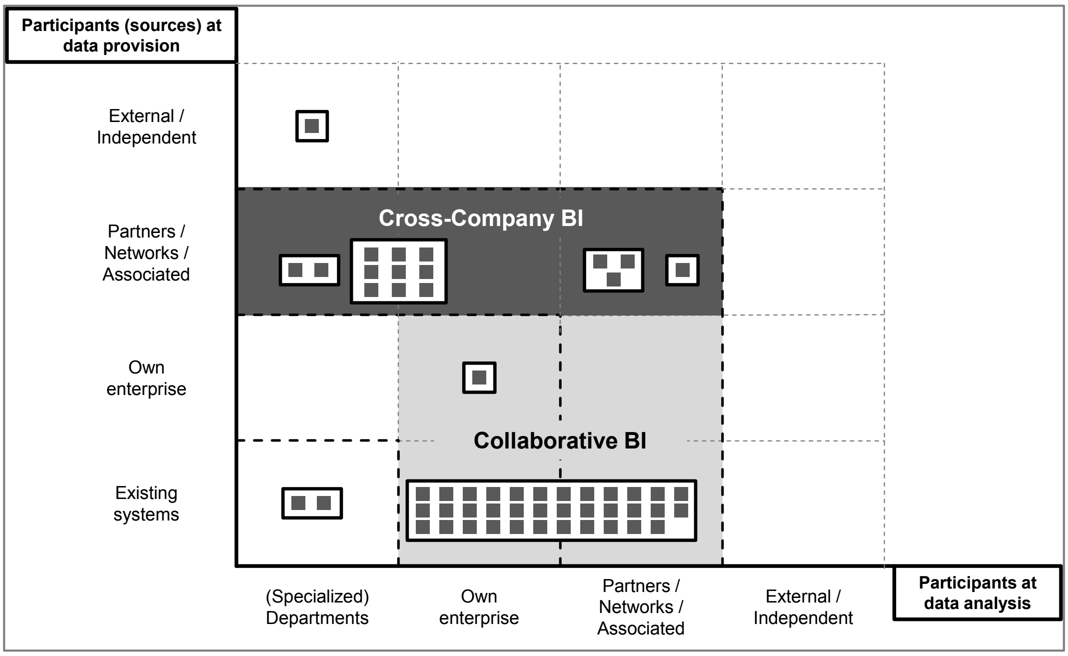


Fig. 2: Orientation framework for CBI and CCBI with the results of the literature reviews

3 A model of CCBI networks

To implement CCBI, network structures between companies have to be formed that accommodate the specific needs of information workers in those partner companies. While different structures are possible, we first present the BIN concept that describes a P2P network of autonomous organizations that share BI functionalities. It focuses on OLAP queries via different nodes in the net and allows for translation chains between not directly coupled peers (Golfarelli et al., 2011, 2012b; Rizzi, 2012). We take this model as a starting point of thought and use the results of the literature review above to conduct an exhaustive search for (partly) related concepts and techniques to show which ideas are well-covered by the model and by research in IS or IT – and which are not or even missing.

3.1 Parallel and Distributed Data Warehouse Concepts

Different DW architectures exist that more or less separate the stored data into different 'data marts', i.e., parts of all data in the physically or virtually integrated DW, and interconnect those data marts. In any case, an integration of the data is necessary, as it is not evaluable otherwise (Sen and Sinha, 2005). An overview of generally suitable approaches for these cases is given by Furtado (2009). Wehrle et al. (2005) describe the concept of creating a distributed DW using a grid of different computers holding the data. All queries are administrated and routed via a central 'coordinator', i.e., a central server system which might itself be spanned over several machines. Even less regulated and completely decentral is the 'brown dwarf', a network of DW nodes that form an overlay based on an underlying P2P network, thus providing redundancy, empirically proven speed improvements, and a working network even if one or more nodes fail (Doka et al., 2011). This does, however, require the allowance of changes in the peers' local systems, which neglects their autonomy. Some work also has been done on OLAP queries in (P2P) network themselves, mainly focusing on strategies for caching results (Kalnis et al., 2002; Kehua and Manirakiza, 2012; Seshadri et al., 2005), chosing the best adjacent peers (Aouiche et al., 2006), and approximating results (Wu et al., 2009). All these approaches provide valuable aspects for a CCBI model, but assume a uniform data scheme, which usually is not the case and cannot always been achieved when different companies or organizations decide to exchange data

(Vaisman et al., 2009). The following paragraph therefore examines the overcoming of different data schemes.

3.2 Dealing with Heterogeneous Multidimensional Data Schemes

Three different basic approaches exist to merge data sets with different multidimensional schemes: (1) Using a central scheme in a central DW for all data, (2) using a distributed ('federated') DW approach with a global scheme, and (3) using autonomous DW systems whose schemes are merged when needed, e.g., based on translation tables (Rizzi, 2012; Tseng and Chen, 2005). As we emphasize the aspect of independence and the possibility to couple or uncouple systems on demand, only the third approach seems reasonably applicable here.

We therefore need to 'match' two or more different data schemes (and data sets) by creating a 'mapping' of the elements, hierarchies, and so on. Very thorough investigations of available matching methods and algorithms have been given by Rahm and Bernstein (2001) and later by Shvaiko and Euzenat (2005). All authors differentiate the matching algorithms on the use of scheme information only or taking other information into account, such as the evaluation of instances, linguistic analyses, or user input. Further work can be found by looking at the concepts of ontology matching (Shvaiko and Euzenat, 2013), as ontologies and their ability to flexibly represent hierarchies are common in a DW context to describe the stored information (Kehlenbeck and Breitner, 2009; Pardillo and Mazon, 2011). All approaches, however, are focused on providing a matching of somewhat flat structures. In a DW, simply finding a mapping on an element-by-element basis might not be sufficient. It cannot be guaranteed that aggregations in dimensions or cubes can be performed properly, e.g., because some elements were mapped to wrong levels, wrong elements, or were not mapped at all (see Fig. 3 for exemplary mapping problems of two differently implemented 'geography' dimensions).

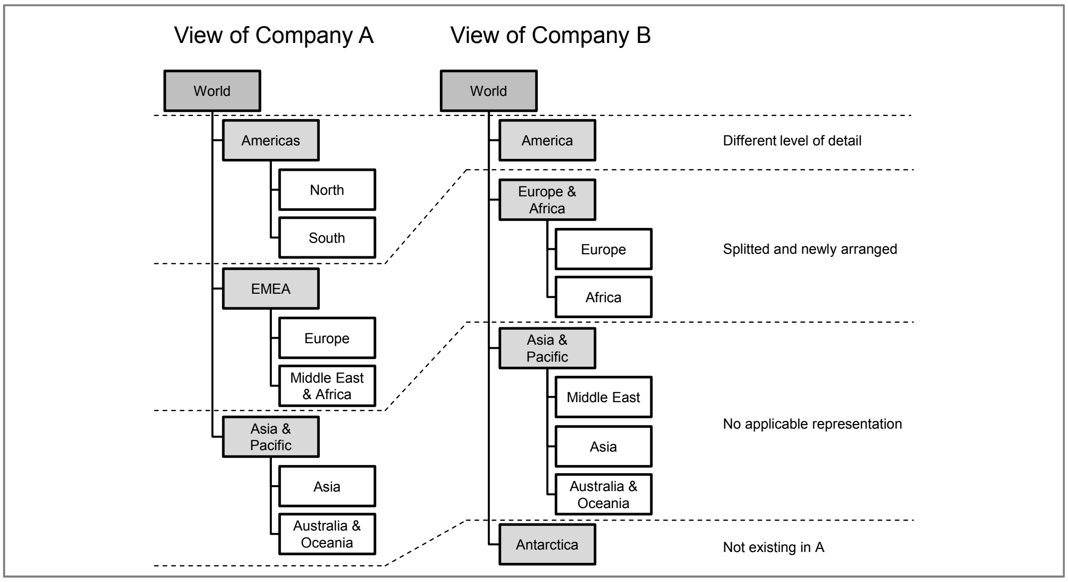


Fig. 3: Mapping problems with different implementations of a 'geography' dimension

The underlying problems are diverse and have been subjects of discussions, but are not yet solved (Berger, 2009; Mazón et al., 2009). This might be the reason, why even the BIN approach addresses the problem, but works with the assumption of a correct mapping (Golfarelli et al., 2012b). While the term ‘correct mapping’ seems to be intuitive with regards to humans defining a mapping of two dimensions and their hierarchies, little research has been done on the systematic, formal description of correctness. Banek et al. (2007) address the problem of creating an algorithm that must preserve the order of dimension levels, but they do not formulate a final condition for that. For the first time Torlone (2008) describes desirable properties of a ‘perfect match’ (considering one dimension on each side) by making use of the order of levels, the completeness of the mapping and the aggregation functions in place. These properties are presented in Tab. 3, using the formal notation introduced in paragraph 2.1 and defining a dimensional matching as a function $\mu: d_1 \rightarrow d_2$, assigning the levels of d_1 to fitting levels of d_2. Riazati et al. (2011) show that enforcing strictness, i.e., the aggregation of elements to one and only one parent, also is a desirable property of dimensions, both for singular use and for mappings.

Tab. 3: Properties of dimension mappings (μ denotes the matching function)

Property	Description
Coherence	After matching, each pair of levels l, l' represents the same hierarchical order as before, i.e., $l \preccurlyeq_1 l' \leftrightarrow \mu(l) \preccurlyeq_2 \mu(l')$.
Soundness	The mapping is complete regarding all elements, i.e., $m_1(l) = m_2(\mu(l))$.
Consistency	The aggregation function ρ is preserved for each pair of levels, i.e., $\rho_1^{l \to l'} = \rho_2^{\mu(l) \to \mu(l')}$.
Strictness	In all dimensions, every aggregation of elements is done with respect to a maximum of one 'parent' element of the superior level, i.e., for each pair of levels l, l', $\rho^{l \to l'}$ is functional, i.e., right-unique.

The given measures present a standard for considerations on defining the goodness of an already finished mapping (Beneventano et al., 2013; Bergamaschi et al., 2011), but do not necessarily present a final set of measures. We therefore use them as a basis for the development of quality measures for mappings, especially in CCBI networks, in chapter 4. We also show that for the mapping of whole cubes they have to be supplemented by other measures that take more information into account, namely the relation of different dimensions to each other (see section 4.3.1).

3.3 Distributed analyses and agents

While distributed data warehouses mainly address the idea of storing and providing data, our former findings have shown that distributed analyses are a part of CCBI, although not often considered in the respective research. One of the most prominent concepts in distributed work is that of intelligent agents, i.e., systems or pieces of software that are able to function autonomously and are able to react, take proactive measures and to function collaboratively including the exchange of information (Wooldridge and Jennings, 1995). The literature review showed that the idea of agents in BI is already established in a one-company setting (Molensky et al., 2010). Bobek and Perko (2006) present a synthesis of agent-based BI research that structures the field in distributed data acquisition, modeling, and delivering. They provide the example of extracting data from different financial organizational sources to build up better scoring models, which fits the intention of CCBI networks. Similar work has been conducted on different data mining algorithms that make use of distributed agents to combine several data sources (Park and Karpguta, 2003), even using P2P network protocols (Giannella et al., 2004). These agents, however, rely on a working network with matched data sources. We therefore focus our considerations on the underlying network and its structure. Once this has been established, moving well-known concepts on agent-based data mining might be an auspicious task (see considerations in chapter 6).

3.4 Security and Trust

Sharing of data and information always leads to concerns about privacy of the participants, particularly in decentralized structures (Kementsietsidis et al., 2003). Some authors deal with the topic specifically addressing OLAP queries in networks (e.g., Agrawal et al. (2005), Cuzzocrea and Bertino (2011, 2014)), some thoroughly discuss security in different levels of P2P networks, e.g., Kirkman and Dezhgosha (2011). Also data mining as an analytic task can make use of distributed structures, but must ensure not to compromise the privacy of the underlying data, e.g., by finding de-anonymizing patterns. Several algorithms for this field have been developed and furthermore improved (Kantarcioglu and Clifton, 2004; Tassa, 2014). These aspects are becoming more and more of interest as the health sector realizes the benefits that arise from sharing information about patients, diseases, and treatments between (independent) medical institutions. Here, security and anonymity become vital for the success of any data sharing initiative, which is why several approaches have been considered for that problem (Kohlmayer et al., 2014; Li et al., 2013; Tassa and Gudes, 2012). The necessary efforts are not less with a centralized approach than with a decentralized one, because institutions permanently transfer data to third parties, which demands for even higher security and privacy for the data (Ermakova and Fabian, 2013). In addition, the concept of trust between the network participants must be considered. Information sharing cannot produce valuable outcome, if some partners withhold information (so called ‘free riders’ (Manshaei et al., 2013)) or even provide false information. All of these problems are known in general networks and gain interest in the BI community (Chang et al., 2006; Ooi et al., 2003). We acknowledge that a comprehensive CCBI network model must take these aspects into account. We save space for them accordingly, but do not elaborate on them here to keep the focus on the functional aspects of the network.

3.5 CroCoBIN: A Simple Reference Model for Peer-to-peer-based CCBI-Networks

The vast amount of existing approaches for different parts of the research topic shows that a lot of aspects have to be considered and have already been discussed, although not necessarily with regards to CCBI networks. When creating a model of these networks, the first decision to be taken is whether the structure should follow a centralistic or decentralistic/distributed approach. As shown in the sections before, centralistic approaches tend to lead to an easier integration of the data, since both technically and semantically the data can be managed at one point using one overlaying data schema. This schema however has to be negotiated upfront. When new partners join the network, it has to be altered or the new partners have to simply adapt to it, regardless of the goodness of fit to their own schemas. The overall discussion on centralization vs. decentralization in IT-based networks has been under revision for decades (King, 1983), and still various results are shown for case studies in both approaches (Malaurent et al., 2012). King (1983) argues that the choice of means mainly depends on the factor of ‘control’, i.e., the power of one or more over the network and the data in it. CCBI focuses on the cooperation of independent partners, which favors an approach where not one company alone can gain control over the data, especially if privacy issues are of high concerns. Additionally, Arigliano et al. (2013) argue that building a comprehensive system implies a substantial risk for SMEs, because for such a network a complex and expensive framework has to be created. We therefore choose to use a P2P network as the underlying structure which will preserve a maximum of autonomy and independence for the partners and allows time-bound couplings. We believe these benefits outweigh the efforts for creating schema mappings in a dynamic environment. As we show in our study, the quality of those networks regarding the available information for all partners actually is comparable to networks with a centralistic schema (see chapter 5). P2P networks also are acknowledged in adjacent research fields, e.g., knowledge management, if centralized structures are uneconomically or impossible to achieve (Maier and Hädrich, 2006). They furthermore allow for the usage of concepts from Semantic P2P (SP2P) networks that provide similar components and ideas, but focus on ontology-based automated data exchange (Mawlood-Yunis et al., 2011). All previous considerations on CCBI networks can be structured by using a very simple first model of such a network as shown in Fig. 4.

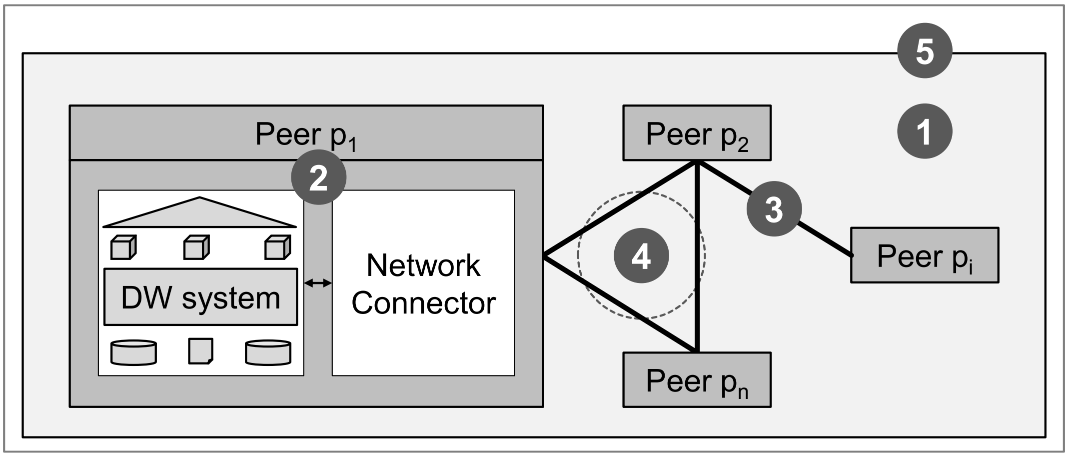


Fig. 4: Five components of a CroCoBIN

A *Cro*ss-*Co*mpany *B*usiness *I*ntelligence *N*etwork (*CroCoBIN*) consists of n peers that hold data in local DW systems and exchange some or all of this data. The five components are:

(1) The *network structure* consists of peers $P = \{p_1, \dots, p_i, \dots, p_n\}$, a function $F_P: P \rightarrow 2^P \backslash \emptyset$, assigning each peer to at least one other, defining its neighborhood $N(p)$ and a metric $s: P \times P \rightarrow \mathbb{R}$, to determine the similarity of two peers‘ data. The net therefore can be represented by a graph $G=(P, E)$ with P being the peers and $E = \bigcup_i N(p_i)$ being all connections (Aberer et al., 2005). We additionally assume that G is connected and not necessarily acyclic and that neighborhood is symmetric, i.e., $p_1 \in N(p_2) \leftrightarrow p_2 \in N(p_1)$, without saying that the *quality* of the connection or translation is identical in both ways.

(2) The *peer structure* consists of the local systems and all parts needed for network activities. A public data store might be separated from a private data store out of security concerns. The public data store must hold the data and the scheme information. The network components include the connector itself and an appropriate query engine. It also can include cache functionalities for internal data, recently provided to others, and external data, recently retrieved or passed through (or even pre-fetched data, which we will not discuss here).

(3) *Processes of peer-pairing* deal with the selection and implementation of matching algorithms. Depending on the public data stores of the peers in each pair, it can be useful to allow different algorithms for different peers, which fully complies with a P2P approach. It is vital, however, to determine the quality of such a connection, to make the pairings comparable to each other and to enable routing strategies, see component (4).

(4) *Processes in the net* describe the overall behavior of peers in the network, especially the routing of queries, i.e., the decision, which peers to connect to, to achieve the best results. Usually this refers to ‘minimal costs’ (Aberer et al., 2005). In a CroCoBIN, costs can be understood as ‘information loss’, when pairing two peers does not result in a perfect mapping and therefore not all information can be transferred, e.g., because one peer has a more detailed data scheme than the other (see chapter 4 for a detailed discussion).

(5) Regulations have to be found by the participants according to their specific business situation. Core regulations might include (a) minimal requirements on local and network hardware, (b) a minimum of connections each peer has to build, so that similarities are possibly high and alternative routes are existing, if participants leave the net, (c) security measures, e.g., encryption, and (d) a minimum set of data that has to provided, so that each peer contributes to the net.

Taking all that into account, a more sophisticated model can be constructed, that serves as a reference model for implementations or theoretical considerations (see Fig. 5). It is abstract enough to incorporate existing approaches, e.g., the BIN or most of the SP2P concept, while simultaneously providing some definite design rules that serve as a common ground for the following developments of quality measures and the findings on CCBI networks built upon them.

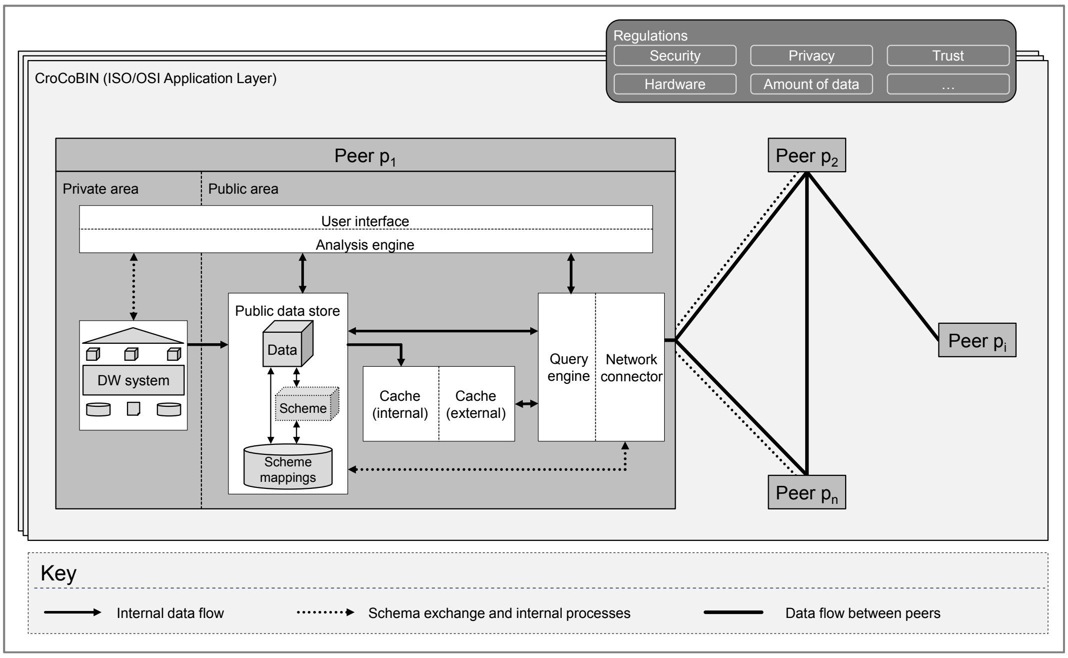


Fig. 5: CroCoBIN Reference Model

4 Quality measurement in CCBI networks

We already argued for the need of quality measurement in CCBI networks. Our demand is supported by the findings of prior work in this area. Melkas (2004) states on inter-company networks: "[…] information-related issues require particularly urgent attention. […] Challenges appear to be especially numerous there. Tools for analyzing information quality in such environments and on the basis of qualitative data have been lacking." (Melkas, 2004, pp. 74, 75) This conforms to the already quoted evaluations of needed research on DW scheme matching and BI networks (Golfarelli et al., 2012b; Riazati et al., 2011; Rizzi, 2012; Torlone, 2008). We therefore structure a CroCoBIN in different levels and explain the respective measurement needs and their relations.

4.1 Notations and Terminology

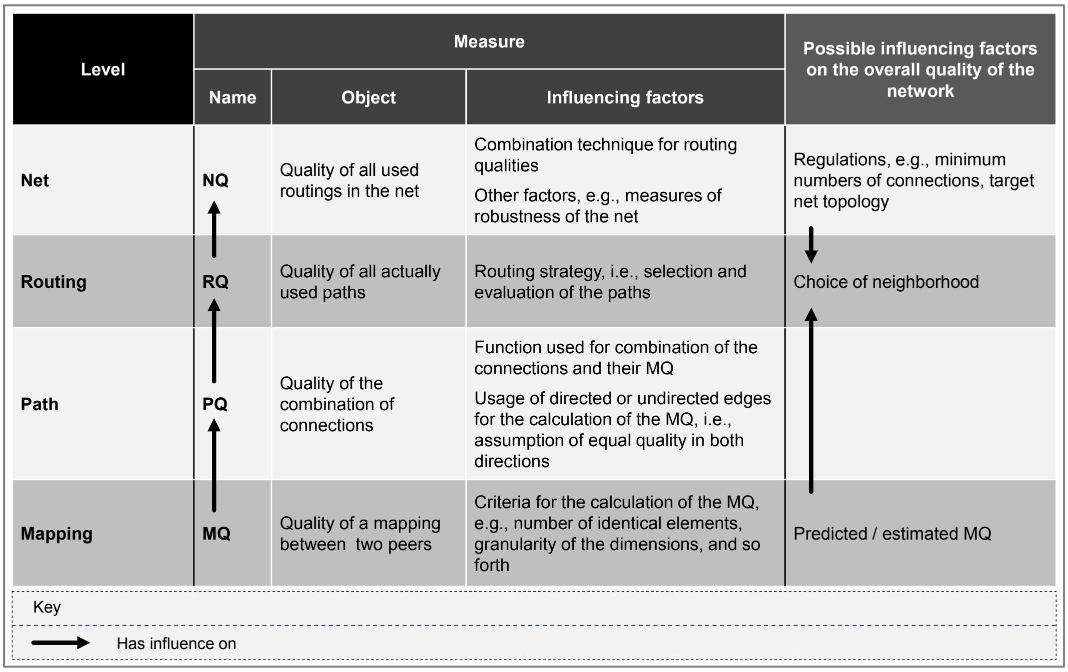
To ease the description of measurement definitions and network parts, we first give a brief introduction into the notation and terminology we use. We have defined a network (graph) G and hereinafter use the standard notation $V = \{v_1, v_2, \ldots, v_n\}$ for the *vertices* (nodes) and $E = \{e_1, e_2, \ldots, e_m\}$ for the *edges* (connections) (Gross and Yellen, 2014). Each edge e has a starting node v_i and an ending node v_j, with $v_i \neq v_j$, and is written $e = (v_i, v_j)$ or simplified $v_i v_j$. In a CroCoBIN each connection is bi-directional, so the graph is undirected. A *path* in a network is an ordered set of vertices $v_0, v_1, \ldots, v_n$. The path's *length* is given by the number of edges in it. A subpath of P is a path consisting only of connected edges of P, holding the same order. $P(v_i, v_j)$ is the path from v_i to v_j.

A node's *degree* $d(v)$ is the number of its neighbors, so the size of its *neighborhood* N. The *density* of a graph $d(G)$ is the relation of all existing to all possible edges, i.e., $d(G) = \frac{2|E|}{|V|(|V|-1)}$. The *distance* of two vertices $d(u, v)$ is the length of the shortest path between them. While the *size* of a graph usually denotes the number of edges in it, we use this term differently to denote the number of vertices in it (which is usually called the graph's *order*). We do so, because in the given business context it seems much more appropriate and adequate to the common understanding of the word, to define the size of a CCBI network by the number of participants instead of the (probably more fluctuating) number of connections.

4.2 Measures in CCBI Networks

The basis of any CCBI network is given by the directly established connections between peers. To rate those connections, we evaluate the quality of the mappings between them. This *mapping quality* (MQ) is defined w.r.t. information loss or similarity and is a descriptive measure that can objectively be evaluated as no direct influence is existing. Data and schemes are predefined and the matching is presumed to create the best possible mapping according to the purpose of multidimensional data exchange. Then, most of the data is exchanged on paths with a length greater one. Thus, an applicable combination is to be developed to calculate the *path quality* (PQ) based on the MQ of the used edges. The selection of the best paths for one node to interact with others leads to a set of paths, the routing. The *routing quality* (RQ) is evaluated accordingly. Finally, the quality of the net as a whole must be measured. As its purpose is to provide a most useful surrounding for the participants, its quality is directly dependent on the information flow in the net. The *net quality* (NQ) therefore combines all RQs in the net, reflecting each peer's gain from participating.

Some influencing factors on the qualities exist. E.g., a minimum or target number of connections for each peer could be defined. An underlying topology, e.g., random vs. scale-free network design, can change the network behavior significantly (Albert et al., 2000; Barabási and Bonabeau, 2003; Seo et al., 2013). This might lead to restrictions in the choice of neighborhood for the peers, which could have a great impact on routing qualities. The choice of neighborhood itself is dependent on the estimated MQ between the peers in question, a factor that itself is interesting for analysis, as it brings uncertainty in practical implementations. In our measure development and studies we take these factors into account as it is shown in the overview on levels, measures, and influencing factors (see Fig. 6).



Level	Measure			Possible influencing factors on the overall quality of the network
	Name	Object	Influencing factors	
Net	**NQ**	Quality of all used routings in the net	Combination technique for routing qualities Other factors, e.g., measures of robustness of the net	Regulations, e.g., minimum numbers of connections, target net topology
Routing	**RQ**	Quality of all actually used paths	Routing strategy, i.e., selection and evaluation of the paths	Choice of neighborhood
Path	**PQ**	Quality of the combination of connections	Function used for combination of the connections and their MQ Usage of directed or undirected edges for the calculation of the MQ, i.e., assumption of equal quality in both directions	
Mapping	**MQ**	Quality of a mapping between two peers	Criteria for the calculation of the MQ, e.g., number of identical elements, granularity of the dimensions, and so forth	Predicted / estimated MQ

Key

→ Has influence on

Fig. 6: Levels, quality measures and influencing factors for CroCoBIN

4.3 Measure and Quality Definitions

A definition of data quality (DQ) has been given with “data that are fit for use by data consumers“ (Wang and Strong, 1996, p. 6). With small variations this publication is considered to be the standard on data quality and (the here interchangeably used term) information quality (Knight and Burn, 2005; Naumann et al., 1999; Scannapieco et al., 2005; Tayi and Ballou, 1998). Scannapieco et al. (2005) recognize four criteria of DQ in a broad state-of-the-art review that are common to all reviewed publications. Data must be *accurate*, i.e., it should not differ from a value that is perceived as a reference. It must be *complete*, as well in quantity as in detail. It must be *current*, i.e., provided shortly after its generation and up-to-date. And finally it must be *consistent*, i.e., conform to specified rules of data integrity. We can see that the currency of the data is nearly impossible to ensure or even to examine when matching peer schemes, but as DWs hold historical data, this problem is somewhat neglectable for our cause. Accuracy is vital, but can only be achieved in the local systems as each peer is solely responsible for its data. Completeness and consistency on the other hand can be measured in a mapping and, as we have seen, Torlone (2008) provides a definition of suitable properties (see Tab. 3). They also can be rearranged to serve as measures, especially for the MQ as shown in the following section, which concludes with a table showing the ‘translation’ of mapping properties to mapping qualities (see Tab. 4).

4.3.1 Mapping Quality (MQ)

Common approaches for schema integration usually consider a ‘master scheme‘, which should be covered by the scheme in question as much as possible. In a P2P network each peer (‘A’) tries to optimize the data received. For MQ measurement therefore A’s scheme is the master scheme and the MQ should be higher, the more data the paired peer (‘B’) can provide for possible fields in the multidimensional data set of A. We call the quality measure for this relation the *true MQ* (TMQ) and define its values to be in the interval [0,1], where 0 means that B does not provide any useful data to A and 1 means that B can provide data for all data fields defined by the scheme of A. This TMQ is only measurable with full knowledge of both schemes and a complete matching process. It is therefore hard to determine in real world scenarios and, as shown in the context of the PQ (section 4.3.2), hardly usable for pre-evaluating possible connection combinations. Still, it provides a solid theoretical control measure that allows testing of a simpler calculated MQ. This prevents the effect of a ‘self-fulfilling prophecy’, when measures are used to value the outcome of algorithms that use the measures themselves (e.g., in Lodi et al. (2008) and Mandreoli et al. (2006)). A simpler MQ is necessary to have a real-world-applicable measure that can be estimated in a short time, so that this *predicted MQ (PMQ)* enables peers to choose the best connections. This is possible, because, as shown, automated schema matchings exist for flat structures and studies show that, compared to matchings by humans, they are able to find about 40 to 80 percent of all possible mappings (Batista and Salgado, 2007; Duchateau and Bellahsene, 2010; Yatskevich et al., 2007).

Therfore we use the (absolute) property of soundness of a dimension mapping (Torlone, 2008) to create a first idea of the MQ, the *MQ_soundness (MQ_s)*, by calculating the amount of mappable elements (of B) compared to the desired maximum (by A) for each dimension:

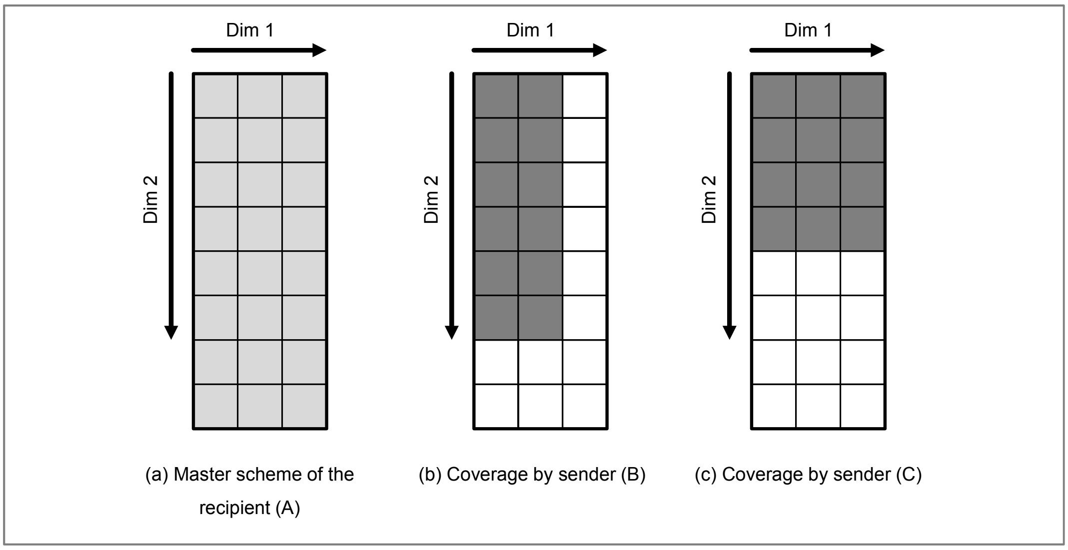
$MQ_s = \frac{1}{|D|}\sum_{d \in D} \frac{\text{No. of mappable elements in } d}{\text{No. of all elements in } d}$, with D being the dimensions in A’s scheme.

So a mapping is completely ‘sound’, i.e., $MQ_s \in [0;1] = 1.0$, if and only if all possible elements in each dimension are given by B. Analogously the *MQ_consistency (MQ$_c$)* can be calculated by looking at the consistency of a mapping, i.e., the coherence of the levels (Torlone, 2008) applied to the elements, which demands even more accuracy:

$MQ_c = \frac{1}{|D|}\sum_{d\in D}\frac{\text{No. of consistently mapped elements in } d}{\text{No. of mapped elements in } d}$, with element m_i in d being consistent, if $\mu(\rho_A(m_i)) = \rho_{\mathrm{B}}(\mu(m_i))$, i.e., the mapping of the parent element of any element (in A) is equal to the parent element (in B) of the mapping of the given element ($MQ_c \in [0;1]$).

Both partial qualities do not sufficiently represent a mapping, though. See Fig. 7 for an example of a two-dimensional data cube (with only one level in each dimension and $3x8 = 24$ items in total), where MQ$_s$ and MQ$_c$ are different for two TMQ-identical mappings ($TMQ(B) = TMQ(C) = \frac{12}{24} = 0{,}5$). Depending on the analysis focus, mapping B could be preferred because of the similar granularity of each dimension, although it is regarded less desirable because of:

$MQ_s(B) = \frac{1}{2}\left(\frac{2}{3}+\frac{6}{8}\right) = \frac{17}{24} < \frac{18}{24} = \frac{1}{2}\left(\frac{3}{3}+\frac{4}{8}\right) = MQ_s(C).$



(a) Master scheme of the recipient (A) (b) Coverage by sender (B) (c) Coverage by sender (C)

Fig. 7: Two exemplary TMQ-identical mappings with different MQ$_s$ and MQ$_c$

To deal with this problem, a third partial quality *MQ_balance (MQ$_b$)* is introduced, that uses the cosine similarity (Nguyen and Bai, 2011; Singhal, 2001) of the vectors that represent the mappings (based on the mappable elements per dimension) and the vector that represents the master scheme (see Fig. 8) to determine the best ‘fit’ to the master scheme.

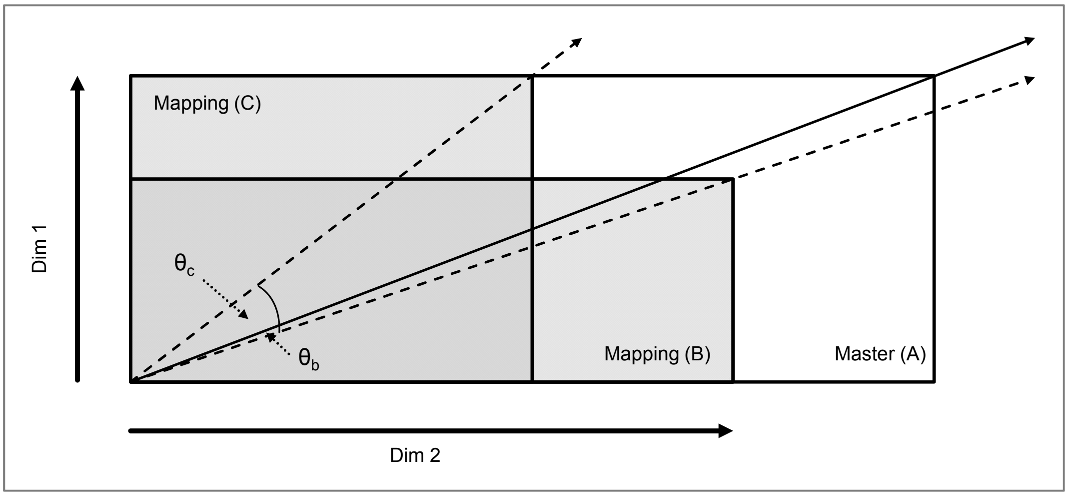


Fig. 8: Vectors of mappings to determine the cosine similarity

Finally, this leads to:

$MQ_b = \cos(\theta)$, with θ being the angle between a and b.

It is $a = \left(|d_1|, \dots, |d_i|, \dots, |d_{|D|}|\right)$, $\mathrm{b} = \left(|\mu(d_1)|, \dots, |\mu(d_i)|, \dots, |\mu(d_{|D|})|\right)$, with $|d_i|$ being the number of elements in dimension i and $|\mu(d_i)|$ being the number of mapped elements for dimension i. $\cos(\theta) = \frac{a \cdot b}{\|a\| \|b\|}$, with $\|a\|$ being the Euclidian norm: $\|a\| = \sqrt{\sum_{i=1}^{n}(a_i)^2}$.

It has been shown that singular measures often are less meaningful than their combination (Rijsbergen, 1979), which has lead to combined and weighted measures in contexts comparable to ours (Aligon et al., 2014; Aouiche et al., 2006; Banek et al., 2007; Duchateau & Bellahsene, 2010; Freire et al., 2012). We therefore propose the final MQ, always to be interpreted from the receiving peer's point of view, i.e., $MQ(v_1, v_2)$ is the quality to be achieved by transferring data from v_2 to v_1:

$MQ = \frac{\alpha MQ_S + \beta MQ_C + \gamma MQ_b}{\alpha + \beta + \gamma}$, with $\alpha, \beta, \gamma \in \mathbb{R}^{\geq 0}$ and $\alpha + \beta + \gamma > 0$.

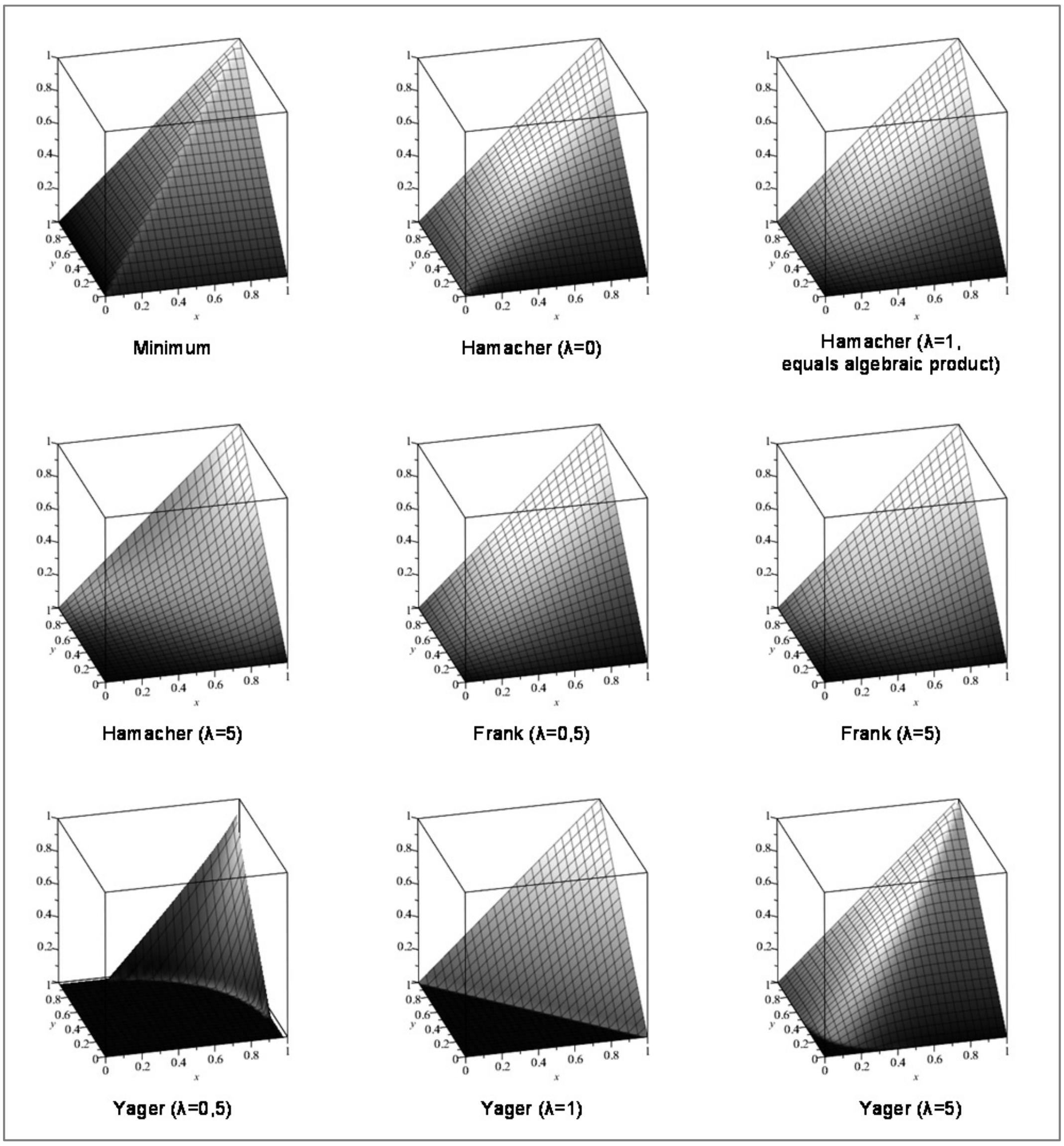
Finally, Tab. 4 presents the relation between the underlying properties of dimension mappings (see Tab. 3) and the partial mapping qualities used.

Tab. 4: Properties of dimension mappings linked to partial mapping qualities

Property	Partial mapping quality	Description
Coherence	**MQ_c**	Coherence and consistency aim to achieve the same thing: The order and aggregation of levels should be the same for mapped dimensions as for the original dimensions. The MQ_c translates this absolute property of a mapping to one that is evaluated for all levels and dimensions individually, so that a gradation of coherence/consistency for a comprehensive mapping is possible.
Soundness	**MQ_s**	Similar to the considerations on coherence, the MQ_s transforms the existence of a complete mapping of all elements to a relative measure which can be evaluated for each dimension individually and then be combined using the average due to its now quantitative rather than categorical nature.
Consistency	**MQ_c**	See coherence.
Strictness		The strictness property (each levels only aggregates to only one other level) is not measured, but assumed as a prerequisite property for participating data warehouses, because otherwise aggregations even in one data warehouse could lead to problems (Riazati et al., 2011).
	MQ_b	The MQ_b has no equivalent in the discussed dimension matching properties, because is does not consider only one dimension, but rather the relation of several dimensions to each other before and after a mapping.

4.3.2 Path Quality (PQ)

The PQ needs to combine different MQ values of the edges on a path from one peer to another, thus describing the ‘information loss’ (more precisely: the information *kept*). The similar concept of ‘semantic degradation’ has been used in Peer Data Management Systems (PDMS) research. Mandreoli et al. (2006) use the algebraic product as a function to combine the MQ values and refer in general to the function families of *t-norms* (with the algebraic product being a special form of the *Hamacher t-norm*). They are suitable because of their useful properties, e.g., being monotone, associative, commutative and providing values according to $t: [0,1] \times [0,1] \rightarrow [0,1]$. Very different t-norms exist and we feel the need to examine (see chapter 5), which are appropriate for the given case as different parametrizations even of the same t-norm family can lead to different rankings of edge combinations (see exemplary t-norm plots in Fig. 9).



Minimum

Hamacher (λ=0)

Hamacher (λ=1, equals algebraic product)

Hamacher (λ=5)

Frank (λ=0,5)

Frank (λ=5)

Yager (λ=0,5)

Yager (λ=1)

Yager (λ=5)

Fig. 9: Plots of nine exemplary t-norms

We therefore only state:

$PQ = t\big(MQ(v_0, v_1), \ldots, MQ(v_i, v_{i+1}), \ldots, MQ(v_{n-1}, v_n)\big)$, with $P(v_0, v_n)$ being the path in question and $t(x_1, \ldots, x_n) = t\left(x_1, t\big(x_2, t(\ldots, x_n)\big)\right)$, t being a t-norm.

Identically to the TMQ a *true PQ (TPQ)* can be defined, that can only be calculated having full knowledge of all schemes and translations in a path, which is highly unlikely in a real world scenario, as all possible paths between all peers would have to be evaluated. TPQ is a $[0,1]$-value, being 1 if all values of the receiving peer in a chain are mapped by values of the sending peer in the same chain (see Fig. 10 for an illustration).

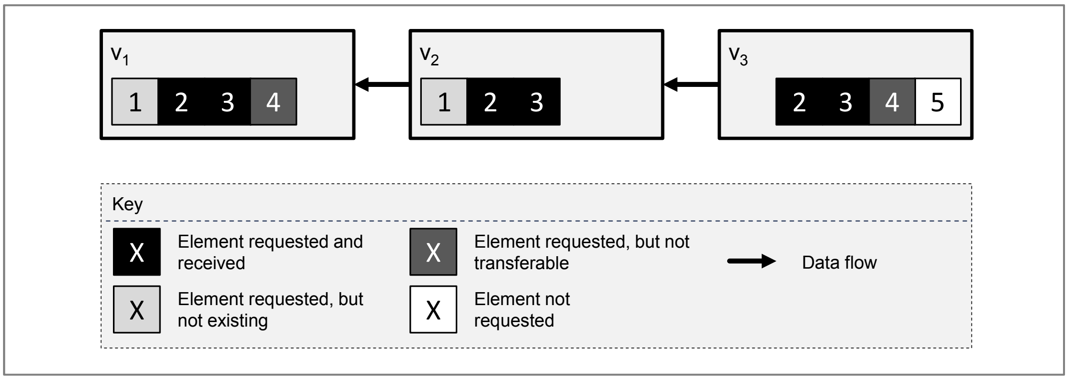


Fig. 10: Illustrative example of a TPQ calculation along a simple path (TPQ=0.5)

4.3.3 Routing Quality (RQ)

A routing represents the set of all used paths by a node (v_1) to all other nodes under the condition, that every node (v_i) is only reachable by exactly one path (P_i), including suppaths of paths from v_1 to any other node v_j, i.e., if $v_i \in P_j$, then P_i is a suppath of P_j. This is to prevent aggregation and summation errors (see section 3.2 and Mazón et al. (2009)). Let G be a new exemplary net with five nodes as shown in Fig. 11.

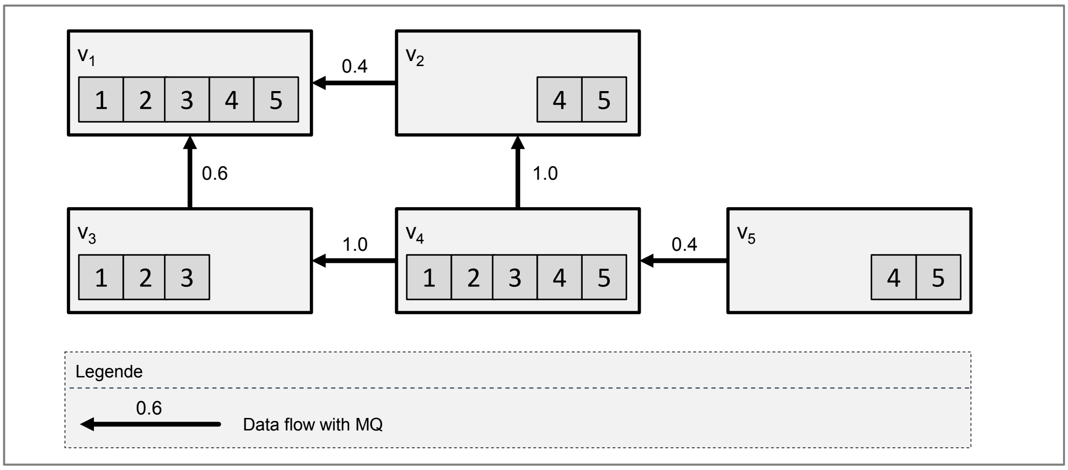


Fig. 11: Alternative routing possibilites in a five-element network

Then looking from v_1, there are different possible routings, e.g. the one using the paths (e_{12}), (e_{13}, e_{34}, e_{45}), and in it (e_{13}, e_{34}) with MQ values 0.6 and 1.0, which seems to be a more promising path from v_1 to v_4 than (e_{12}, e_{24}) with MQ values 0.4 and 1.0. However, simple counting of the elements that can be received by this routing (8) shows, that the alternative routing (e_{13}), (e_{12}, e_{24}, e_{45}), leads to more overall transferred elements (9), because v_3 cannot map the elements of v_5. Therefore, a static routing strategy (Medhi and Ramasamy, 2007) that takes all paths into account is to be used. Although the network should be (and is) dynamic, fluctuation and net size are small enough to efficiently recreate a static routing table regularly. The routing quality for a node v_1 then can be simply derived by the average path quality of all used paths:

$RQ = \frac{1}{n-1}\sum_{i=2}^{n} PQ(v_1, v_i)$, with n being the size of the net.

The *true RQ (TRQ)* is calculated analogously to the TPQ as the mean TPQ of all used paths.

4.3.4 Net Quality (NQ)

All previous considerations only deal with the quality of the mappings. On the network level, other properties may be considered. They include the robustness of the net, i.e., the ability to function, when nodes or connections are lost. Network robustness is a complex field of study and a lot of different measures exist (Albert and Barabási, 2002; Chen et al., 2012; Grubesic et al., 2008; Jamakovic and Uhlig, 2007; Manzano et al., 2011). To avoid blurring of data quality measurement with not directly comparable values, we renounce mixing those values into the net quality and only use the *average path length (AvgPL)* and the *clustering coefficient (CC)*, that quantifies the connection strength of a neighborhood (Albert & Barabási, 2002; Watts and Strogatz, 1998) as supplemental descriptive values. The net quality therefore simply is the mean of the RQ values and expresses the amount of data that is kept despite of information loss in the mappings in the net. It is controlled by the *true NQ (TNQ)*, calculated like the other 'true' qualities:

$NQ = \frac{1}{n}\sum_{i=1}^{n} RQ(v_i)$, with $v_i \in V$ and n being the size of the net.

5 Network simulation and results

Now that quality measures exist, we put them to use by simulating a CroCoBIN with varying network and mapping parameters to determine whether the qualities provide useful results (by testing them against the true qualities) and to generate recommendations on how to structure CroCoBINs in the real world, e.g., which net sizes are probably useable.

5.1 Simulation Model and Implementation Design

We created a prototypical implementation of a network simulation tool (CroCoSIM) based on C# that is able to depict the specific mappings between multidimensional DW nodes and to calculate all presented measures. It allows parametrization of network topologies, network sizes, target degrees of connections for each node, strategies for neighborhood selection, MQ estimation for the mapping, weighting of the partial qualities of the MQ and the combination function (t-norm). In addition it uses pseudo-random values that ensure identical preconditions in various parametrizations. A configurable seed then guarantees different basic values to prevent misleading results of a single data set. The peer's schemes are based on a three-dimensional master scheme that is pseudo-randomly reduced according to different parameters, so that a minimal mapping of one (top) element is guaranteed and mappings can be created automatically. Note, that more complex schemes would only lead to more complex mappings, the range of values for the MQ, however, would not change in principle. The net is generated using a basic algorithm for scale-free nets (the 'BA algorithm' (Albert & Barabási, 2002)) or random nets similar to Nobari et al. (2011), but instead of reducing a full graph, we build it up from a connections-free graph ('modified random algorithm' (MRA)). Both variants are modified to use the PMQ for each mapping, so that connections are not made randomly with just any node. See Tab. 5 and Tab. 6 for a parameter description.

Tab. 5: Parameters for net generation in CroCoSIM

Parameter	Description
Basic algorithm	Creates a random or scale-free network based on a modified algorithm to incorporate the predicted MQ.
Net size	Defines the target size of the network. $net\_size \in \mathbb{N}^{>0}$
Target degree	Random nets only. Describes the target average(!) degree of every node, thus defining the density of the net. $target\_degree \in \mathbb{N}^{>1}$
m	Scale-free nets only. Describes the number of nodes, every new node connects to. Implicitly describes the size m_0 of a fully connected core network. $m \in \mathbb{N}^{>0}, m_0 = m + 1$ Densities of scale-free (G) and random (H) networks are comparable with $d(G_{2m}) \cong d(H_{target\_degree})$.
Core net size	Random nets only. Describe the size of a fully connected core network, analogous to scale-free nets' m_0. $corenet\_size \in \mathbb{N}^{>0}$
Seed	The seed of the pseudo-random generator for the net. Seed of other pseudo-random generators, e.g., for the dimension and cube creation, are functional dependent.

Tab. 6: Parameters for neighborhood selection in CroCoSIM

	Applicable to		
Parameter	MRA	BA	Description
PMQ goodness (PMQG)	X	X	Describes the precision of the estimation/prediction of the MQ. 0 equals random numbers, 1 equals a perfect prediction. $PMQG \in [0; 1]$
Target degree spread	X		Describes the possibility to deviate the actual degree of any node from the target degree. 0 equals a forced target degree for each node. $target\_degree\_spread \in [0; 1]$
Sorted node list		X	Checks peers for potential peering in a sorted manner. Increases the chance to chose a node with a high PMQ. $sorted \in \{true; false\}$
μ		X	Describes the effect of the PMQ on the 'attractivity' of the existing nodes in the net (Chen and Shi, 2004). 0 equals the ignoring of the PMQ, 1 changes the node's attractivity to the one of the node with the highest degree. $\mu \in \mathbb{R}^{\geq 0}$, Overall attractivity$\sim \frac{\mu \times PMQ \times \max(d) + d}{1+\mu}$
Selfish approach	X	X	If set to 'false', pairing selection will be based on the provided(!) MQ to other nodes rather than on the usually used received MQ. $selfish \in \{true; false\}$
MQ (α, β, γ)	X	X	Describes the weights of the partial MQ values MQ_s (α), MQ_c (β) and MQ_b (γ).

5.2 Preliminary Studies

In our simulation, we followed the dependencies of the qualities in the network. Therefore, we first identified a good parameter combination of α, β, and γ for the partial qualities of the MQ. Using a regression analysis between the TMQ and the partial qualities of the MQ on more than one million generated pairings, we identified fitting values $\alpha = 0.5836, \beta = 0.0556$, and $\gamma = 0.3608$. These values are not directly applicable, because MQ and TMQ do not have a linear relation. To overcome this issue, we tested 231 combinations of α, β, and γ, where we only used 0.05-steps of the parameters with $\alpha+\beta+\gamma=1$, checked the correlation using Spearman's rho and found the combination $\alpha = 0.60, \beta = 0.05$ und $\gamma = 0.35$ to be best-performing, which nearly equals the regression values and thus is used for all former simulations.

As a next step, we analyzed a broad variety of different network combinations, generating nearly 16 million paths, which were used to compare the PQ values (and the accompanying RQ and NQ values) with their respective true qualities (TPQ, TNQ, TRQ) using 35 different t-norms from three different t-norm families (Hamacher, Yager, and Frank; see Tab. 9 and Tab. 10 in the appendix for detailed definitions and results). The goal was to identify a t-norm that would lead to a PQ that most accurately reflects the TPQ and besides that leads to high TPQ values, providing a most effective network in terms of information exchange. Using four net size classes that would usually occur in real-world cases of SME networks ($n = 25, 50, 100, 200$), we found the t-norm $T^Y_{1.25}$, i.e., the Yager-norm with λ=1.25, to be the single best norm that would reach the given goals in all network sizes. It is therefore used in all network simulations.

As a reference for all considerations regarding the TNQ, we created completely connected networks for all parameter combinations, i.e., every node had a direct mapping to every other node. In our given setting, the reference values were size-dependent as expected, reaching a maximum of 0.1171 with $n = 200$ ($n = 25\colon 0.0951, n = 50\colon 0.1003, n = 100\colon 0.1111$).

5.3 Network parameter analysis

We conducted a first preliminary network study to find out which parameters have the greatest effect on network quality by conducting a ceteris-paribus (cp) analysis. In descending order the most important factors are net size, average degree, PMQG, basic algorithm, the choice of a selfish approach, and lastly the seed. Based on the results and on additional considerations, e.g. defining a PMQG range of 0.4 to 0.8 (because that is roughly what automated matching can achieve, see section 4.3.1) which highly reduced the cp-influence of it, we set values for the second study (see Tab. 7).

Tab. 7: Parameter values of different network studies (@: interval step size)

Parameter	1st study values	2nd study values	3rd study values
Basic algorithm	Random, Scale-free	Random, Scale-free	Random
Net size	25,50,100,200	25,50,100,200	25,50,100,200
Target degree	[4,10] @ 1	[4,10] @ 2	[4,10] @ 2
m	[2,5] @ 1	[2,5] @ 1	
Core net size	1, 5	1	1
Seed	[1,3] @ 1	[1,6] @ 1	[1,6] @ 1
PMQG	[0.0,1.0] @ 0.1	[0.4,0.8] @ 0.1	[0.4,0.8] @ 0.1
Target degree spread	[0.0,1.0] @ 0.1	[0.1,1.0] @ 0.1	[0.1,1.0] @ 0.1
Sorted node list	True, False	True	
μ	[0.00, 5.00] @ 0.25	[0.00, 2.25] @ 0.25	
Selfish approach	True, False	True, False	True

We then could see that NQ and TNQ values of networks are highly dependent on the different parameters and that random networks with a ‘selfish’ approach are proven to be superior to scale-free networks and ‘non-selfish’ networks in every size class ($p < 0.005$). See Fig. 12 for a graphical impression of the network quality distribution and Tab. 11 in the appendix for the average values.

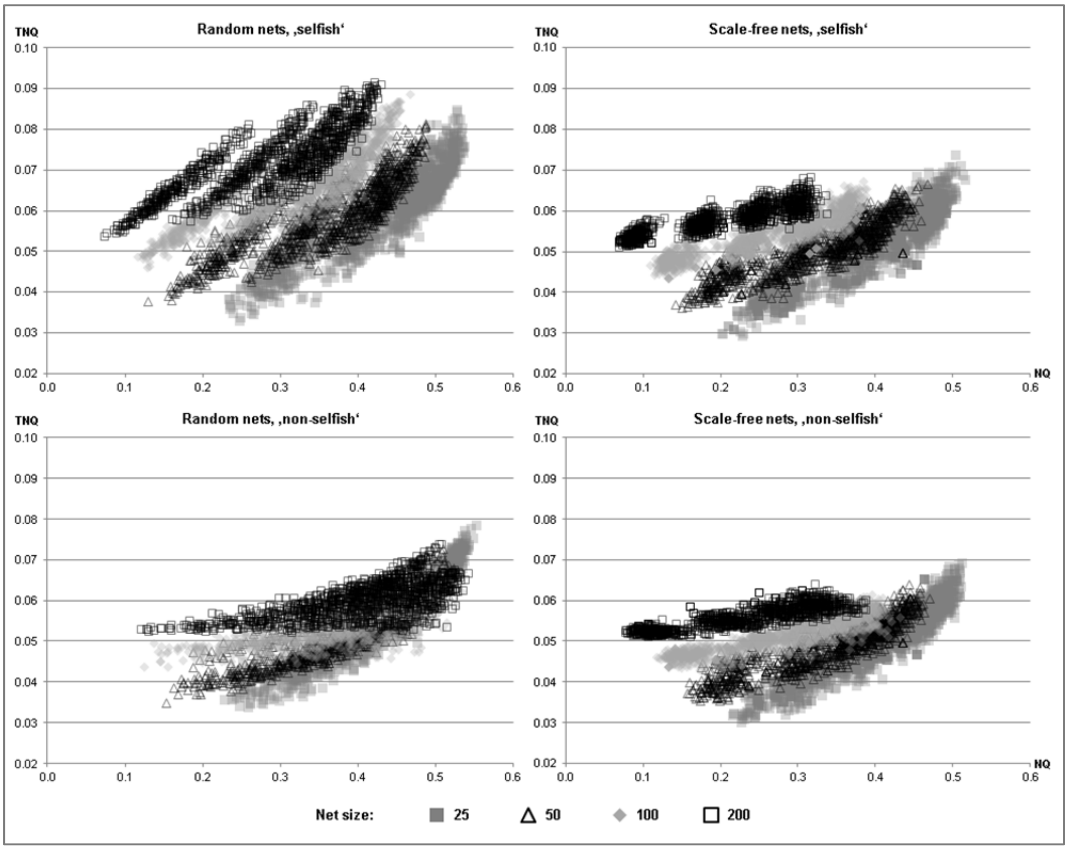


Fig. 12: NQ and TNQ values of different network types

A third study then was conducted, focusing only on the 'best' networks (see Tab. 7). The graphical analysis suggests that bigger networks lead to higher TNQ values. Regarding to the mean value and with no respect to the different average degrees this is significantly true for networks of the sizes 50, 100, and 200 ($p < 0.001$), but cannot be proven for $n = 50$ in comparison to $n = 25$ ($p = 0.2339$). It can be explained by the relatively high connection count of 'small' networks. Once this effect does not take place any longer, because the net has reached a relevant size, bigger nets lead to higher *true* qualities. That is remarkable, as it suggests that company networks will actually profit from taking new partners into the net. This phenomenon is similar to the well-known 'critical mass effect' which can be observed in many networks, e.g., for telecommunication (Katz & Shapiro, 1994; Oren and Smith, 1981). We confirmed this behavior in a supplementary fourth study to take effect for nets up to the size of ca. 300 participants, before path lengths increased so much, that information loss was overweighting the effect. Note, that the average TNQ value for the nets is very good, e.g. for $n = 200$ and an average degree of 7 it is at 0.0722 of a maximum of 0.1171 (average degree of 199(!)) with some networks reaching values over 0.9.

Finally, Tab. 8 shows the influence of the parameters (third study values) and their interaction terms (Wooldridge, 2013) in a regression analysis on NQ and TNQ values.

Tab. 8: Influence of parameters on NQ and TNQ

Parameter	Influence on NQ	Influence on TNQ
Degree	0.9476	0.6196
Net size	- 0.8186	0.6495
Net size x PMQG	0.2339	0.1201
PMQG	0.2028	0.2094
Spread	0.1835	- 0.0421
Degree x Spread	- 0.1777	- 0.0120
Degree x PMQG	- 0.1577	0.0753
Spread x PMQG	0.0874	0.3855
Net size x Spread	0.0798	0.0783
Net size x Degree	0.0615	- 0.3449

As assumed before, net size and degree are the most prominent variables. Note, that a bigger net size increases the TNQ, but decreases the NQ as path lengths take great effect on the t-norm combinations of MQ values. It therefore is vital to formulate further hypotheses with reference to size classes. Besides that, we can see that not forcing an equal connection count for each peer, i.e., allowing a high degree spread, improves NQ and TNQ for higher values of PMQ, because nodes that serve as 'hubs' can be created. Thus it is recommended not to set too strict regulations when designing a cross-company network.

5.4 Forecasting Abilities of Simple Qualities

Using the calculated values and different (partly multi-level) regression models, we created estimation functions on MQ and NQ values to forecast the TMQ and TNQ that could be expected to be existing in networks. Those are

$$\widehat{TMQ} = 5.7161 \times MQ^3 - 7.600 \times MQ^2 + 3.3930 \times MQ + 0.5026,$$

with $R^2 = 0.9934$ and

$$\widehat{TNQ} = \left(0{,}0241 \times \log_2 \frac{n}{25} + 0{,}0786 \times \log_2 \frac{d}{4} + 0{,}1866\right) \times NQ + \left(0{,}0239 \times \log_2 \frac{n}{25} - 0{,}0467 \times \log_2 \frac{d}{4} - 0{,}0140\right),$$

d being the average degree in the net, with $R^2 = 0.9219$.

As both estimated qualities show a very high coefficient of determination, the quality measures MQ and NQ that participants of the network can calculate on their own are beneficial for network creation and probably lead to high true network qualities.

6 Conclusions

Cross-company BI has been a rising topic of discussion over the past few years. We considered existing work about collaborative BI and different similar fields to provide a working definition. On this basis we developed CroCoBIN, a reference model that provides a solid base for research in many ways, of which we pursued the one to practical quality measurement. We have shown that CCBI networks are able to work on a P2P basis and that already a small degree of connections between the peers brings up nets that are comparative to networks with a central scheme, because their information upkeep can be at most as high as the one in fully connected networks. P2P networks provide for independence of the participants and allow time-bound connections, which is extremely useful for companies that form networks based on projects or do not have the wish or capability to form long-lasting relations with dominant, probably less flexible data structures. Furthermore, we showed that increasing the number of participants might even lead to more powerful nets, thus encouraging companies to share data and to invite others.

Although conducted on a very extensive simulation basis, our study has limits. First and foremost, it was done on real-world-alike data, but not on real data. Our results show that conducting a field study might pay off for the participants and we highly recommend such a case study. Second, although we particularly mention the ability of P2P networks to adapt to changing environments and constant (un-)coupling of peers as a benefit of such a network design, our studies could be considered as 'static' simulations. We did not focus on the dynamics that arise when the element of time is made a more important issue in the use of the networks. The data warehouses used by the participants could be used only for a short period of time while they were beneficial, e.g., when covering data about current events, natural catastrophes, special fairs, etc. Further studies can focus on this element more and see if the network configurations uphold their identified properties in a more dynamic and time-bound environment – especially when the efforts of creating mappings is accounted for more precisely. It also will be interesting to discover how (maybe agent-based) distributed analyses can be implemented in these networks with their structure-changing nature.

The reference model and our search for covered topics show a lot of other options for further research. Those include the identification of suitable matching algorithms, defining appropriate cache technologies, and evaluating different regulations for CCBI networks, especially focusing on security and trust issues. The quality measures we developed will then provide a comparable measurement of the quality to be achieved in very different network configurations.

References

Aberer, K., Alima, L. O., Ghodsi, A., Girdzijauskas, S., Haridi, S., & Hauswirth, M. (2005). The essence of P2PA a reference architecture for overlay networks. In G. Caronni (Ed.), *Proceedings of the Fifth IEEE International Conference on Peer-to-Peer Computing* (pp. 11–20). Los Alamitos, CA: IEEE.

Agrawal, R., Srikant, R., & Thomas, D. (2005). Privacy preserving OLAP. In F. Ozcan (Ed.), *Proceedings of the 2005 ACM SIGMOD international conference on Management of data (SIGMOD '05)* (pp. 251–262). New York, NY: ACM.

AIS. (2011). *Senior Scholars' Basket of Journals*. Retrieved from http://start.aisnet.org/?SeniorScholarBasket

Albert, R., & Barabási, A.-L. (2002). Statistical mechanics of complex networks. *Reviews of Modern Physics, 74*(1), 47–97.

Albert, R., Jeong, H., & Barabasi, A.-L. (2000). Error and attack tolerance of complex networks. *Nature, 406*(6794), 378–382. Retrieved from http://dx.doi.org/10.1038/35019019

Ale Ebrahim, N., Ahmed, S., & Taha, Z. (2009). Virtual R&D Teams in Small and Medium Enterprises: A Literature Review. *Scientific Research and Essays, 4*(13), 1575–1590.

Aligon, J., Golfarelli, M., Marcel, P., Rizzi, S., & Turricchia, E. (2014). Similarity measures for OLAP sessions. *Knowledge and Information Systems, 39*(2), 463–489. doi:10.1007/s10115-013-0614-1

Alwis, D. (2004). The role of intellectual capital in organisational value creation: An application of a theoretical model to two case studies. In Brunel University London (Ed.), *Dept of Computer Science Theses* . London: Brunel University London. Retrieved from http://bura.brunel.ac.uk/handle/2438/5068

Aouiche, K., Jouve, P.-E., & Darmont, J. (2006). Clustering-Based Materialized View Selection in Data Warehouses. In Y. Manolopoulos, J. Pokorný, & T. Sellis (Eds.), *Advances in Databases and Information Systems* (pp. 81–95). Berlin, Heidelberg, New York: Springer.

Arigliano, F., Azzini, A., Braghin, C., Caforio, A., Ceravolo, P., Damiani, E., . . . Zavatarelli, F. (2013). Knowledge and Business Intelligence Technologies in Cross-Enterprise Environments for Italian Advanced Mechanical Industry. In *Knowledge and Business Intelligence Technologies in Cross-Enterprise Environments for Italian Advanced Mechanical Industry* (pp. 104–110). Citeseer.

Baars, H., Felden, C., Gluchowski, P., Hilbert, A., Kemper, H.-G., & Olbrich, S. (2014). Shaping the Next Incarnation of Business Intelligence. *Business & Information Systems Engineering, 6*(1), 11–16.

Baars, H., & Kemper, H.-G. (2010). Business Intelligence in the Cloud? In *Proceedings of the 2010 Pacific Asia Conference on Information Systems (PACIS)* (pp. 1528–1539).

Banek, M., Vrdoljak, B., Tjoa, A. M., & Skocir, Z. (2007). Automating the Schema Matching Process for Heterogeneous Data Warehouses. In I. Y. Song, J. Eder, & T. M. Nguyen (Eds.): *Vol. 4654. Lecture Notes in Computer Science, Proceedings of the 9th International Conference on Data Warehousing and Knowledge Discovery (DAWAK)* (pp. 45–54). Berlin, Heidelberg: Springer.

Barabási, A.-L., & Bonabeau, E. (2003). Scale-Free Networks. *Scientific American, 288*(5), 50–59.

Barkai, D. (2001). *Peer-to-peer computing: Technologies for sharing and collaborating on the net*. Hillsboro, OR: Intel Press.

Batista, M. d. C. M., & Salgado, A. C. (2007). Information Quality Measurement in Data Integration Schemas. In C. Koch, J. Gehrke, M. N. Garofalakis, D. Srivastava, K. Aberer, A. Deshpande, . . . C.-C. Kanne (Eds.), *Proceedings of the 33rd international conference on very large data bases (VLDB 2007)* (pp. 61–72). Wien: ACM.

Beneventano, D., Olaru, M. O., & Vincini, M. (2013). Analyzing Dimension Mappings and Properties in Data Warehouse Integration. In R. Meersman, H. Panetto, T. Dillon, J. Eder, Z. Bellahsene, N. Ritter, . . . D. Dou (Eds.), *On the Move to Meaningful Internet Systems: OTM 2013 Conferences: Confederated International Conferences: CoopIS, DOA-Trusted Cloud, and ODBASE 2013, Graz, Austria, September 9-13, 2013. Proceedings* (pp. 616–623). Berlin, Heidelberg: Springer Berlin Heidelberg.

Bergamaschi, S., Olaru, M.-O., Sorrentino, S., & Vincini, M. (2011). Semi-automatic Discovery of Mappings Between Heterogeneous Data Warehouse Dimensions. *International Journal on Information Technology, 1*(3), 38–46. Retrieved from http://doi.searchdl.org/01.IJIT.1.3.15

Berger, S. (2009). *FedDW: A model-driven approach for querying federations of autonomous data marts*. PhD Thesis. Linz, Austria: Johannes Kepler Universität.

Bloom, N., Garicano, L., Sadun, R., & van Reenen, J. (2014). The Distinct Effects of Information Technology and Communication Technology on Firm Organization. *Management Science, 60*(12), 2859–2885. doi:10.1287/mnsc.2014.2013

Bobek, S., & Perko, I. (2006). Intelligent Agent based Business Intelligence. In A. Méndez-Vilas, A. Solano Martín, J. A. Mesa González, & J. Mesa González (Eds.), *Current Developments in Technology-Assisted Education* (pp. 1047–1051). Badajoz, Spain: FORMATEX.

Böhringer, M., Gluchowski, P., Kurze, C., & Schieder, C. (2010). A Business Intelligence Perspective on the Future Internet. In Association for Information Systems (Ed.), *Proceedings of the 16th American Conference on Information Systems (AMCIS)* . Red Hook, NY: Curran.

Bramer, W. M., Giustini, D., & Kramer, B. M. R. (2016). Comparing the coverage, recall, and precision of searches for 120 systematic reviews in Embase, MEDLINE, and Google Scholar: A prospective study. *Systematic Reviews, 5*(1), 7. doi:10.1186/s13643-016-0215-7

Cabibbo, L., & Torlone, R. (1998). A Logical Approach to Multidimensional Databases. In H.-J. Schek, G. Alonso, F. Saltor, & I. Ramos (Eds.), *Advances in Database Technology EDBT'98* (pp. 183–197). Berlin, Heidelberg: Springer.

Chang, E., Dillon, T. S., & Hussain, F. K. (2006). *Trust and reputation for service oriented environments: Technologies for building business intelligence and consumer confidence*. Chichester: Wiley.

Chen, Q., & Shi, D. (2004). The modeling of scale-free networks. *Physica A: Statistical Mechanics and its Applications, 335*(1–2), 240–248. doi:10.1016/j.physa.2003.12.014

Chen, X., Li, H., & Ke, M. (2012). Study on Complex Network Stability Based on the Structure Parameters. In Association for Information Systems (Ed.), *Proceedings of the 2012 International Conference on Information Systems (ICIS)* (pp. 521–526). Red Hook, NY: Curran.

Cuzzocrea, A., & Bertino, E. (2014). A Comprehensive Theoretical Framework for Privacy Preserving Distributed OLAP. In R. Meersman, H. Panetto, A. Mishra, R. Valencia-García, A. Soares, I. Ciuciu, . . . H. Chan (Eds.), *Lecture Notes in Computer Science. On the Move to Meaningful Internet Systems: OTM 2014 Workshops* (pp. 117–136). Berlin, Heidelberg: Springer.

Cuzzocrea, A., & Bertino, E. (2011). Privacy Preserving OLAP over Distributed XML Data. *Journal of Computer and System Sciences, 77,* 965–987.

Dayal, U., Vennelakanti, R., Sharma, R., Castellanos, M., Hao, M., & Patel, C. (2008). Collaborative Business Intelligence: Enabling Collaborative Decision Making in Enterprises. In Z. Tari & R. Meersmann (Eds.), *On the Move to Meaningful Internet Systems: OTM 2008* (pp. 8–25). Berlin, Heidelberg: Springer.

DeSanctis, G., & Gallupe, R. B. (1987). A Foundation for the Study of Group Decision Support Systems. *Management Science, 33*(5), 589–609. doi:10.1287/mnsc.33.5.589

Doka, K., Tsoumakos, D., & Koziris, N. (2011). Brown Dwarf: A fully-distributed, fault-tolerant data warehousing system. *Journal of Parallel and Distributed Computing, 71*(11), 1434–1446. doi:10.1016/j.jpdc.2011.07.008

Duchateau, F., & Bellahsene, Z. (2010). Measuring the Quality of an Integrated Schema. In J. Parsons, M. Saeki, P. Shoval, C. Woo, & Y. Wand (Eds.), *Lecture Notes in Computer Science. Conceptual Modeling – ER 2010* (pp. 261–273). Berlin, Heidelberg: Springer.

Ermakova, T., & Fabian, B. (2013). Secret Sharing for Health Data in Multi-provider Clouds. In *Proceedings of the 2013 IEEE 15th Conference on Business Informatics* (pp. 93–100). IEEE Computer Society.

Fayyad, U. M., Piatetsky-Shapiro, G., & Smyth, P. (1996). Advances in knowledge discovery and data mining. In U. M. Fayyad, G. Piatetsky-Shapiro, P. Smyth, & R. Uthurusamy (Eds.), *Advances in knowledge discovery and data mining* (pp. 1–34). Menlo Park et al.: AAAI Press et al.

Forbes. (2015). *Roundup Of Analytics, Big Data & Business Intelligence Forecasts And Market Estimates, 2015*. Retrieved from http://www.forbes.com/sites/louiscolumbus/2015/05/25/roundup-of-analytics-big-data-business-intelligence-forecasts-and-market-estimates-2015/#1690e33b4869

Freire, C., Souza, B. F. F., Souza, D., Batista, M. C. M., & Salgado, A. C. (2012). Towards an information quality approach to enhance query routing processes. In E. Pardede (Ed.), *Proceedings of the 14th International Conference on Information Integration and Web-based Applications & Services (IIWAS '12)* (pp. 311–314). New York, NY: ACM.

Furtado, P. (2009). A Survey of Parallel and Distributed Data Warehouses. *International Journal of Data Warehousing and Mining, 5*(2), 57–77. doi:10.4018/jdwm.2009040103

Gartner Inc. (2016). *Gartner Says Worldwide Business Intelligence and Analytics Market to Reach $16.9 Billion in 2016*. Retrieved from http://www.gartner.com/newsroom/id/3198917

Gartner Inc. (2014). *Hype Cycle for Business Intelligence and Analytics, 2014, G00262070*. Retrieved from https://www.gartner.com/doc/2813223/hype-cycle-business-intelligence-analytics

Gartner Inc. (2013). *Gartner Says Organizations Must Treat Information as an Asset in its Own Right*. Retrieved from http://www.gartner.com/newsroom/id/2506315

Gartner Inc. (2011). *Hype Cycle for Business Intelligence, 2011, G00216086*. Retrieved from https://www.gartner.com/doc/1766215/hype-cycle-business-intelligence-

Giannella, C., Bhargava, R., & Kargupta, H. (2004). Multi-agent Systems and Distributed Data Mining. In M. Klusch, S. Ossowski, V. Kashyap, & R. Unland (Eds.), *Cooperative Information Agents VIII: 8th International Workshop, CIA 2004, Erfurt, Germany, September 27-29, 2004. Proceedings* (pp. 1–15). Berlin, Heidelberg: Springer Berlin Heidelberg.

Gill, T. G., & Hevner, A. R. (2013). A Fitness-Utility Model for Design Science Research: Article 5. *ACM Transaction on Management Information Systems, 4*(2), 5.

Golfarelli, M., Mandreoli, F., Penzo, W., Rizzi, S., & Turricchia, E. (2012a). A Query Reformulation Framework for P2P OLAP. In Nicola Ferro & Letizia Tanca (Eds.), *Proceedings of the 20th Italian Symposium on Advanced Database Systems (SEBD)* (pp. 147–154). Padova: Edizioni Libreria Progetto.

Golfarelli, M., Mandreoli, F., Penzo, W., Rizzi, S., & Turricchia, E. (2012b). OLAP query reformulation in peer-to-peer data warehousing. *Information Systems, 37*(5), 393–411.

Golfarelli, M., Mandreoli, F., Penzo, W., Rizzi, S., & Turricchia, E. (2011). BIN: Business intelligence networks. In M. E. Zorrilla, J.-N. Mazón, Ó. Ferrández, I. Garrigós, F. Daniel, & J. Trujillo (Eds.), *Business Intelligence Applications and the Web* (pp. 244–265). Hershey, PA: IGI Global.

Golfarelli, M., Mandreoli, F., Penzo, W., Rizzi, S., & Turricchia, E. (2010). Towards OLAP query reformulation in peer-to-peer data warehousing. In I.-Y. Song & C. Ordonez (Eds.), *Proceedings of the 13th ACM International Workshop on Data Warehousing and OLAP (DOLAP)* (pp. 37–44). New York, NY: ACM.

Gross, J. L., & Yellen, J. (2014). Fundamentals of Graph Theory. In J. L. Gross, J. Yellen, & P. Zhang (Eds.), *Discrete mathematics and its applications. Handbook of graph theory* (2nd ed., pp. 2–20). Boca Raton, FL: CRC Press.

Grubesic, T. H., Matisziw, T. C., Murray, A. T., & Snediker, D. (2008). Comparative Approaches for Assessing Network Vulnerability. *International Regional Science Review, 31*(1), 88–112. doi:10.1177/0160017607308679

Harzing, A.-W. (2012). A preliminary test of Google Scholar as a source for citation data: a longitudinal study of Nobel prize winners. *Scientometrics, 94*(3), 1057–1075. doi:10.1007/s11192-012-0777-7

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design Science in Information Systems Research. *MIS Quarterly, 28*(1), 75–105.

Inmon, W. H. (2005). *Building the Data Warehouse* (4th ed.). Indianapolis: Wiley.

Jamakovic, A., & Uhlig, S. (2007). Influence of the network structure on robustness. In IEEE (Ed.), *15th IEEE International Conference on Networks, 2007 (ICON 2007)* (pp. 278–283). Adelaide, Australia: IEEE.

Kalnis, P., Ng, W. S., Ooi, B. C., Papadias, D., & Tan, K.-L. (2002). An adaptive peer-to-peer network for distributed caching of OLAP results. In M. J. Franklin, B. Moon, & A. Ailamaki (Eds.), *Proceedings of the 2002 ACM SIGMOD International Conference on Management of Data* (pp. 25–36). New York, NY: ACM.

Kantarcioglu, M., & Clifton, C. (2004). Privacy-Preserving Distributed Mining of Association Rules on Horizontally Partitioned Data. *IEEE Trans. on Knowl. and Data Eng., 16*(9), 1026–1037. doi:10.1109/TKDE.2004.45

Katz, M. L., & Shapiro, C. (1994). Systems Competition and Network Effects. *The Journal of Economic Perspectives, 8*(2), 93–115. doi:10.2307/2138538

Kaufmann, J. (2015). *Gestaltung unternehmensübergreifender Business-Intelligence-Netzwerke. Dissertation*. Duisburg, Essen: DuEPublico.

Kaufmann, J., & Chamoni, P. (2014). Structuring Collaborative Business Intelligence: A Literature Review. In IEEE (Ed.), *Proceedings of the 47th Hawaii International Conference on System Sciences (HICSS '47)* (pp. 3738–3747). Los Alamitos, CA: IEEE.

Kehlenbeck, M., & Breitner, M. (2009). Ontology-Based Exchange and Immediate Application of Business Calculation Definitions for Online Analytical Processing. In T. B. Pedersen, M. K. Mohania, & A. M. Tjoa (Eds.), *Lecture Notes in Computer Science. Data Warehousing and Knowledge Discovery* (pp. 298–311). Berlin, Heidelberg: Springer.

Kehua, Y., & Manirakiza, A. (2012). Efficient and Semantic OLAP Aggregation Queries in a Peer to Peer Network. *International Journal of Information and Electronics Engineering, 2*(5), 697–701. doi:10.7763/IJIEE.2012.V2.190

Kementsietsidis, A., Arenas, M., & Miller, R. J. (2003). Mapping data in peer-to-peer systems. In Y. Papakonstantinou, A. Halevy, & Z. Ives (Eds.), *Proceedings of the 2003 ACM SIGMOD international conference on Management of data* (pp. 325–336). New York, NY: ACM.

King, J. L. (1983). Centralized versus decentralized computing: organizational considerations and management options. *ACM Comput. Surv., 15*(4), 319–349. doi:10.1145/289.290

Kirkman, S. S., & Dezhgosha, K. (2011). Security review of P2P applications and networks. In H. R. Arabnia, V. A. Clincy, & Solo, Ashu M. G. (Eds.), *Proceedings of the 2011 International Conference on Internet Computing, ICOMP 2011* . O.O.: CSREA Press.

Klarmann, A., Roth, M., & Franczyk, B. (2013). Entscheidungsunterstützung in logistischen Netzwerken - Ansätze für eine Umsetzung. In H. Baars (Ed.): *Vol. 1048. CEUR Workshop Proceedings, WSBI-2012. Workshop Business Intelligence 2012* (pp. 28–36). Stuttgart: CEUR-WS.

Knight, S.-A., & Burn, J. (2005). Developing a Framework for Assessing Information Quality on the World Wide Web. *Informing Science Journal, 8*(5), 159–172.

Kohlmayer, F., Prasser, F., Eckert, C., & Kuhn, K. A. (2014). A flexible approach to distributed data anonymization. *Special Issue on Informatics Methods in Medical Privacy, 50,* 62–76. doi:10.1016/j.jbi.2013.12.002

Laudon, K. C., & Laudon, J. P. (2006). *Management information systems* (9th ed.). Upper Saddle River, NJ: Pearson Prentice Hall.

Li, M., Yu, S., Zheng, Y., Ren, K., & Lou, W. (2013). Scalable and Secure Sharing of Personal Health Records in Cloud Computing Using Attribute-Based Encryption. *IEEE Trans. Parallel Distrib. Syst., 24*(1), 131–143. doi:10.1109/TPDS.2012.97

Liu, L., & Daniels, H. (2012). Towards a Value Model for Collaborative, Business Intelligence-supported Risk Assessment. In *Proceedings of the 6th International Workshop on Value Modeling and Business Ontology (VMBO 2012)* (pp. 1–5). Retrieved from http://vmbo2012.isis.tuwien.ac.at/?page_id=4

Lodi, S., Penzo, W., Mandreoli, F., Martoglia, R., & Sassatelli, S. (2008). Semantic peer, here are the neighbors you want! In A. Kemper (Ed.), *Proceedings of the 11th international conference on Extending database technology (EDBT '08)* (pp. 26–37). New York, NY: ACM.

Maier, R., & Hädrich, T. (2006). Centralized versus peer-to-peer knowledge management systems. *Knowledge and Process Management, 13*(1), 47–61.

Malaurent, J., Yan, L., & Avison, D. E. (2012). Reopening The Centralization-Decentralization Debate: A Comparative Case Study Of ERP Implementation In Two Chinese Petroleum Companies. In *PACIS 2012 Proceedings* (p. 74).

Mandreoli, F., Martoglia, R., Sassatelli, S., & Penzo, W. (2006). SRI: exploiting semantic information for effective query routing in a PDMS. In A. Bonifati & I. Fundulaki (Eds.), *Proceedings of the 8th ACM International Workshop on Web Information and Data Management (WIDM 2006)* (pp. 19–26). New York, NY: ACM.

Manshaei, M. H., Zhu, Q., Alpcan, T., Bacşar, T., & Hubaux, J.-P. (2013). Game theory meets network security and privacy. *ACM Computing Surveys, 45*(3), 1–39. doi:10.1145/2480741.2480742

Manzano, M., Calle, E., & Harle, D. (2011). Quantitative and qualitative network robustness analysis under different multiple failure scenarios. In IEEE (Ed.), *3rd International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT)* . Budapest, Hungary: IEEE.

Markus, M. L., & Silver, M. S. (2008). A Foundation for the Study of IT Effects: A New Look at DeSanctis and Poole’s Concepts of Structural Features and Spirit. *Journal of the Association for Information Systems, 9*(10), 609–633.

Martins, V. A., daCosta, J. P. L., & deSousa, R. T. (2012). Architecture of a Collaborative Business Intelligence Environment based on an Ontology Repository and Distributed Data Services. In INSTICC (Ed.), *Proceedings of the 4th International Conference on Knowledge Management and Information Sharing (KMIS 2012)* (pp. 99–106). Barcelona: INSTICC.

Mawlood-Yunis, A.-R., Weiss, M., & Santoro, N. (2011). A Reference Model for Semantic Peer-to-Peer Networks. In S. Spaccapietra (Ed.), *Lecture Notes in Computer Science. Journal on Data Semantics XV* (pp. 37–65). Berlin, Heidelberg: Springer.

Mazón, J.-N., Lechtenbörger, J., & Trujillo, J. (2009). A survey on summarizability issues in multidimensional modeling. *Data & Knowledge Engineering, 68*(12), 1452–1469. doi:10.1016/j.datak.2009.07.010

Medhi, D., & Ramasamy, K. (2007). *Network routing. The Morgan Kaufmann series in networking*. Amsterdam et al.: Elsevier; Morgan Kaufmann.

Meier, J. J., & Conkling, T. W. (2008). Google Scholar’s Coverage of the Engineering Literature: An empirical study. *The Journal of Academic Librarianship, 34*(3), 196–201.

Melkas, H. (2004). Analyzing Information Quality In Virtual Service Networks With Qualitative Interview Data. In M.I.T. Information Quality Program (Ed.), *Proceedings of the 9th International Conference on Information Quality (ICIQ-04)* (pp. 74–88). Cambridge, MA: Massachusetts Institute of Technology.

Mentzer, J. T., DeWitt, W., Keebler, J. S., Min, S., Nix, N. W., Smith, C. D., & Zacharia, Z. G. (2001). Defining Supply Chain Management. *Journal of Business Logistics, 22*(2), 1–25. doi:10.1002/j.2158-1592.2001.tb00001.x

Mettler, T., & Raber, D. (2011). Developing a collaborative business intelligence system for improving delivery reliability in business networks. In K.-D. Thoben, V. Stich, & A. Imtiaz (Eds.), *Proceedings of the 17th International Conference on Concurrent Enterprising (ICE'11)* (pp. 1–7). Aachen: FIR e. V. at RWTH Aachen University.

Miller, M. (2001). *Discovering P2P*. San Francisco: Sybex.

Molensky, L., Ketter, W., Collins, J., Bloemhof-Ruwaard, J. M., & Koppel, H. van de (2010). Business intelligence gap analysis: A user, supplier and academic perspective. In T. Bui & M. Jarke (Eds.), *Proceedings of the 12th International Conference on Electronic Commerce (ICEC '10)* (Vol. 133, pp. 119–128). New York, NY: ACM.

Muntean, M. (2012). Business Intelligence Approaches. *Mathematical Models & Methods in Applied Sciences, 1,* 192–196.

Naumann, F., Leser, U., & Freytag, J. C. (1999). Quality-driven Integration of Heterogenous Information Systems. In M. P. Atkinson, M. E. Orlowska, P. Valduriez, S. B. Zdonik, & M. L. Brodie (Eds.), *Proceedings of the 25th International Conference on Very Large Data Bases (VLDB '99)* (pp. 447–458). San Francisco, CA: Morgan Kaufmann.

Nguyen, H., & Bai, L. (2011). Cosine Similarity Metric Learning for Face Verification. In R. Kimmel, R. Klette, & A. Sugimoto (Eds.), *Lecture Notes in Computer Science. Computer Vision – ACCV 2010* (pp. 709–720). Berlin, Heidelberg: Springer.

Nobari, S., Lu, X., Karras, P., & Bressan, S. (2011). Fast random graph generation. In A. Ailamaki, S. Amer-Yahia, J. Pate, T. Risch, P. Senellart, & J. Stoyanovich (Eds.), *Proceedings of the 14th International Conference on Extending Database Technology (EDBT/ICDT '11)* (pp. 331–342). New York, NY, USA: ACM.

Ooi, B., Liau, C., & Tan, K.-L. (2003). Managing Trust in Peer-to-Peer Systems Using Reputation-Based Techniques. In G. Dong, C. Tang, & W. Wang (Eds.), *Lecture Notes in Computer Science. Advances in Web-Age Information Management* (pp. 2–12). Berlin, Heidelberg: Springer.

Oren, S. S., & Smith, S. A. (1981). Critical Mass and Tariff Structure in Electronic Communications Markets. *The Bell Journal of Economics, 12*(2), 467–487. doi:10.2307/3003567

Pardillo, J., & Mazon, J. N. (2011). Using Ontologies for the Design of Data Warehouses. *International Journal of Database Management Systems, 3*(2), 73–87. doi:10.5121/ijdms.2011.3205

Park, B., & Karpguta, H. (2003). Distributed Data Mining: Algorithms, Systems, and Applications. In N. Ye (Ed.), *The handbook of data mining* (pp. 341–358). Lawrence Erlbaum Associates.

Power, D. J. (2009). *A Brief History of Decision Support Systems, version 4.1*. Retrieved from http://dssresources.com/history/dsshistory.html

Rahm, E., & Bernstein, P. A. (2001). A Survey of Approaches to Automatic Schema Matching. *The VLDB Journal, 10*(4), 334–350. doi:10.1007/s007780100057

Riazati, D., Thom, J. A., & Zhang, X. (2011). Enforcing strictness in integration of dimensions: beyond instance matching. In I.-Y. Song, A. Cuzzocrea, & K. C. Davis (Eds.), *Proceedings of the 14th ACM International Workshop on Data Warehousing and OLAP (DOLAP)* (pp. 9–16). New York, NY: ACM.

Rijsbergen, C. J. van. (1979). *Information retrieval* (2nd ed.). London, Boston: Butterworths.

Rizzi, S. (2012). Collaborative Business Intelligence. In M.-A. Aufaure & E. Zimányi (Eds.), *Business Intelligence* (pp. 186–205). Berlin, Heidelberg: Springer.

Rodriguez, C., Daniel, F., Casati, F., & Cappiello, C. (2010). Toward Uncertain Business Intelligence: The Case of Key Indicators. *IEEE Internet Computing, 14*(4), 32–40. doi:10.1109/MIC.2010.59

Rothwell, R., & Dodgson, M. (1991). External linkages and innovation in small and medium-sized enterprises. *R & D Management, 21*(2), 125–138.

Scannapieco, M., Missier, P., & Batini, C. (2005). Data Quality at a Glance. *Datenbank-Spektrum, 14,* 6–14.

Schoder, D., & Fischbach, K. (2003). Peer-to-peer prospects. *Communications of the ACM, 46*(2), 27. doi:10.1145/606272.606294

Schoder, D., & Fischbach, K. (2002). Peer-to-Peer. *WIRTSCHAFTSINFORMATIK, 44*(6), 587–589. doi:10.1007/BF03250877

Scholta, C. (2005). *Erfolgsfaktoren unternehmensübergreifender Kooperation am Beispiel der mittelständischen Automobilzulieferindustrie in Sachsen*. Chemnitz: Technische Universität Chemnitz.

Schwalm, S., & Bange, C. (2004). Einsatzpotenziale von XML in Business-Intelligence-Systemen. *WIRTSCHAFTSINFORMATIK, 46*(1), 5–14. doi:10.1007/BF03250991

Sen, A., & Sinha, A. P. (2005). A comparison of data warehousing methodologies. *Communications of the ACM, 48*(3), 79–84. doi:10.1145/1047671.1047673

Seo, H., Kim, W., Lee, J., & Youn, B. (2013). Network-based approaches for anticancer therapy (Review). *International journal of oncology, 43*(6), 1737–1744. doi:10.3892/ijo.2013.2114

Seshadri, S., Cooper, B. F., & Liu, L. (2005). *CubeCache: Efficient and Scalable Processing of OLAP Aggregation Queries in a Peer-to-Peer Network. Technical Report GIT-CERCS-07-12*. Retrieved from http://hdl.handle.net/1853/20072

Sharda, R., Delen, D., & Turban, E. (2014). *Business intelligence and analytics* (10th ed.). Boston, MA: Pearson.

Shvaiko, P., & Euzenat, J. (2013). Ontology Matching: State of the Art and Future Challenges. *IEEE Transactions on Knowledge and Data Engineering, 25*(1), 158–176. doi:10.1109/TKDE.2011.253

Shvaiko, P., & Euzenat, J. (2005). A Survey of Schema-Based Matching Approaches. In S. Spaccapietra (Ed.), *Lecture Notes in Computer Science. Journal on Data Semantics IV* (pp. 146–171). Berlin, Heidelberg: Springer.

Simon, A. R., & Shaffer, S. L. (2002). Chapter 7 - Core Technologies and Building Blocks. In A. R. Simon & S. L. Shaffer (Eds.), *Data Warehousing And Business Intelligence For e-Commerce* (pp. 125–171). San Francisco: Morgan Kaufmann.

Singhal, A. (2001). Modern Information Retrieval: A Brief Overview. *Bulletin of the IEEE Computer Society Technical Committee on Data Engineering, 24*(4), 35–42.

Tassa, T. (2014). Secure Mining of Association Rules in Horizontally Distributed Databases. *IEEE Transactions on Knowledge and Data Engineering, 26*(4), 970–983.

Tassa, T., & Gudes, E. (2012). Secure distributed computation of anonymized views of shared databases. *ACM Trans. Database Syst., 37*(2), 1–43. doi:10.1145/2188349.2188353

Tayi, G. K., & Ballou, D. P. (1998). Examining Data Quality. *Communications of the ACM, 41*(2), 54–57. doi:10.1145/269012.269021

Torlone, R. (2008). Two approaches to the integration of heterogeneous data warehouses. *Distributed and Parallel Databases, 23*(1), 69–97. doi:10.1007/s10619-007-7022-z

Tseng, F. S. C., & Chen, C.-W. (2005). Integrating heterogeneous data warehouses using XML technologies. *Journal of Information Science, 31*(3), 209–229.

Turban, E., Liang, T.-P., & Wu, S. (2011). A framework for adopting collaboration 2.0 tools for virtual group decision making. *Group decision and negotiation, 20*(2), 137–154.

Vaisman, A. A., Minuto Espil, M., & Paradela, M. (2009). P2P OLAP: Data model, implementation and case study. *Information Systems, 34*(2), 231–257. doi:10.1016/j.is.2008.07.001

Vera-Baquero, A., Colomo-Palacios, R., & Molloy, O. (2013). Business Process Analytics Using a Big Data Approach. *IT Professional, 15*(6), 29–35. doi:10.1109/MITP.2013.60

Wang, R. Y., & Strong, D. M. (1996). Beyond Accuracy: What Data Quality Means to Data Consumers. *Journal of Management Information Systems, 12*(4), 5–33.

Watson, H. J., & Wixom, B. H. (2007). The Current State of Business Intelligence. *Computer, 40*(9), 96–99. doi:10.1109/MC.2007.331

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature, 393*(6684), 440–442. doi:10.1038/30918

Webster, J., & Watson, R. T. (2002). Analyzing the Past to Prepare for the Future: Writing a Literature Review. *MIS Quarterly, 26*(2), xiii–xxiii.

Wehrle, P., Miquel, M., & Tchounikine, A. (2005). A model for distributing and querying a data warehouse on a computing grid. In IEEE (Ed.), *Proceedings of 11th International Conference on Parallel and Distributed Systems (ICPADS)* (pp. 203–209). Los Alamitos, CA: IEEE.

Werner, A., Hrach, C., Heyden, K., Alt, R., & Franczyk, B. (2010). Value Chain Cockpit auf Basis betrieblicher Anwendungssysteme. In M. Schumann, L. M. Kolbe, M. H. Breitner, & A. Frerichs (Eds.), *Proceedings der Multikonferenz Wirtschaftsinformatik 2010 (MKWI 2010)* (pp. 1583–1595). Göttingen: Universitätsverlag Göttingen.

WKWI. (2008). WI-Orientierungslisten. *WIRTSCHAFTSINFORMATIK, 50*(2), 155–163.

Wooldridge, J. M. (2013). *Introductory econometrics* (5th ed.). Mason, OH: South-Western Cengage Learning.

Wooldridge, M., & Jennings, N. R. (1995). Intelligent Agents: Theory and Practice. *Knowledge Engineering Review, 10,* 115–152.

Wu, S., Jiang, S., Ooi, B. C., & Tan, K.-L. (2009). Distributed online aggregations. *Proceedings of the VLDB Endowment, 2*(1), 443–454. doi:10.14778/1687627.1687678

Wuertz, G., Ardilio, A., Lasi, H., & Warschat, J. (2013). Towards mass individualization: Life-cycle oriented configuration of time-variable product-service systems. In *2013 Proceedings of PICMET '13: Technology Management in the IT-Driven Services (PICMET)* .

Yatskevich, M., Giunchiglia, F., McNeill ,Fiona, & Shvaiko, P. (2007). *OpenKnowledge Deliverable 3.3: A methodology for ontology matching quality evaluation* (Technical Report DIT-07-062). University of Trento, Trento.

Appendix

Tab. 9: Names and definitions of t-norm families

Name	Definition	
Minimum	$T_M(x,y) = \min(x,y)$	
Hamacher	$T_\lambda^H(x,y) = \begin{cases} 0 \\ \frac{xy}{\lambda+(1-\lambda)(x+y-xy)} \end{cases}$	, for $\lambda = x = y = 0$, for $\lambda \in]0;\infty[$
Frank	$T_\lambda^F(x,y) = \begin{cases} T_M(x,y) \\ T_1^H(x,y) \\ \log_\lambda\left(1 + \frac{(\lambda^x - 1)(\lambda^y - 1)}{\lambda - 1}\right) \end{cases}$	, for $\lambda = 0$, for $\lambda = 1$, for $\lambda \in]0;1[\mathrm{U}]1;\infty[$
Yager	$T_\lambda^Y(x,y) = \max(1 - \left((1-x)^\lambda + (1-x)^\lambda\right)^{\frac{1}{\lambda}}, 0)$	, for $\lambda \in]0;\infty[$

Tab. 10: Ranks and z-values for different t-norms (Hamacher(0.0) is the algebraic product, Frank(0.0) is the minimum-norm)

	Rank Correlation	Rank Mean TNQ	Rank Complete	z-Value Complete	Rank $n = 25$	Rank $n = 50$	Rank $n = 100$	Rank $n = 200$
1	Yag(5.000)	Ham(5.00)	Yag(1.25)	0.3904	6	3	2	1
2	Yag(10.00)	Ham(10.00)	Yag(1.00)	0.3863	3	2	26	24
3	Fra(0.00)	Ham(3.00)	Ham(5.00)	0.3219	4	4	27	25
4	Yag(3.00)	Fra(10.00)	Ham(10.00)	0.3088	1	1	28	27
5	Yag(2.00)	Yag(1.00)	Fra(10.00)	0.3043	7	6	22	15
6	Ham(0.00)	Ham(2.00)	Ham(3.00)	0.2975	5	5	25	19
7	Yag(1.75)	Fra(5.00)	Ham(2.00)	0.2924	8	7	23	21
8	Yag(1.25)	Ham(1.75)	Fra(1.50)	0.2914	15	15	15	3
9	Ham(0.25)	Fra(3.00)	Fra(3.00)	0.2899	11	9	19	11
10	Yag(1.50)	Ham(1.50)	Fra(2.00)	0.2889	12	12	17	18
11	Yag(1.00)	Fra(2.00)	Fra(1.75)	0.2877	14	13	14	14
12	Fra(0.25)	Yag(1.25)	Fra(5.00)	0.2876	9	8	20	17
13	Ham(0.50)	Fra(1.75)	Ham(1.50)	0.2870	13	10	18	16
14	Fra(0.50)	Ham(1.25)	Yag(1.50)	0.2867	26	16	1	2
15	Ham(0.75)	Fra(1.50)	Fra(1.25)	0.2864	17	17	13	4
16	Fra(0.75)	Fra(1.25)	Ham(1.75)	0.2863	10	11	21	13
17	Ham(1.00)	Ham(1.00)	Ham(1.25)	0.2838	16	14	16	12
18	Fra(1.25)	Fra(0.75)	Fra(0.75)	0.2801	19	19	11	6
19	Fra(1.50)	Ham(0.75)	Ham(1.00)	0.2796	20	18	12	5
20	Fra(1.75)	Fra(0.50)	Ham(0.75)	0.2737	21	20	10	8
21	Ham(1.25)	Yag(0.75)	Fra(0.50)	0.2711	22	22	9	7
22	Fra(2.00)	Ham(0.50)	Ham(0.50)	0.2622	24	23	8	9
23	Fra(3.00)	Yag(1.50)	Ham(0.00)	0.2554	18	26	3	23
24	Ham(1.50)	Fra(0.25)	Fra(0.25)	0.2534	25	24	7	10
25	Fra(10.00)	Ham(0.25)	Ham(0.25)	0.2529	23	25	6	22
26	Ham(5.00)	Ham(0.00)	Yag(1.75)	0.1996	27	27	5	20
27	Ham(2.00)	Yag(1.75)	Yag(0.75)	0.1606	2	21	30	29
28	Ham(1.75)	Yag(0.50)	Yag(2.00)	0.1124	29	28	4	26
29	Fra(5.00)	Yag(2.00)	Yag(3.00)	-0.1466	30	29	24	28
30	Ham(3.00)	Yag(0.25)	Yag(5.00)	-0.4801	32	31	29	30
31	Ham(10.00)	Yag(0.00)	Yag(0.50)	-0.8481	31	30	33	32
32	Yag(0.75)	Yag(3.00)	Yag(10.00)	-0.9365	33	32	31	31
33	Yag(0.50)	Yag(5.00)	Fra(0.00)	-1.2322	34	33	32	33
34	Yag(0.25)	Yag(10.00)	Yag(0.25)	-2.0619	28	34	34	34
35	Yag(0.00)	Fra(0.00)	Yag(0.00)	-2.0729	35	35	35	35

Tab. 11: NQ and TNQ values for different networks (2nd study)

Net size	Basic algorithm	‚selfish‘	NQ	TNQ
25	Random	true	0.4282	0.0592
		false	0.4426	0.0559
	Scale-free	true	0.3958	0.0507
		false	0.4042	0.0494
50	Random	true	0.3547	0.0589
		false	0.3890	0.0529
	Scale-free	true	0.3221	0.0505
		false	0.3340	0.0485
100	Random	true	0.3161	0.0659
		false	0.3957	0.0558
	Scale-free	true	0.2639	0.0539
		false	0.2877	0.0516
200	Random	true	0.2816	0.0722
		false	0.4033	0.0602
	Scale-free	true	0.2031	0.0581
		false	0.2334	0.0561

Diskussionsbeiträge der Fakultät für Betriebswirtschaftslehre
Mercator School of Management (MSM)
Universität Duisburg-Essen, Campus Duisburg

---

404. J. KAUFMANN
Cross-Company Business Intelligence: Definition, Model, and Quality Measurement, August 2016.

403. A. KIMMS, I. KOZELETSKYI
Consideration of Multiple Objectives in Horizontal Cooperation with an Application to Transportation Planning, Juni 2016.

2015

402. M. MAIWALD, A. KIMMS
Bi-Objective Safe and Resilient Urban Evacuation Planning, Dezember 2015.

401. M. MAIWALD, A. KIMMS
An Exact Network Flow Formulation for Cell-Based Evacuation in Urban Areas, November 2015.

400. G. SCHMITZ, J. LERCH
Kundenseitige Nutzungsprozesse von Mobilitätsleistungen: Konzeptionelle Grundlagen und empirische Befunde, August 2015.

399. K. SEEKIRCHER, A. KIMMS
Network Design to Anticipate Selfish Evacuation Routing, Juni 2015.

398. D. WEIMAR
Sportindustrie als ökonomische Testumgebung, Januar 2015.

2014

397. A. KIMMS, I. KOZELETSKYI, A. MECA MARTÍNEZ
Bi-Allocation Games: A new class of NTU Games, Dezember 2014.

396. A. ELIAS, A. KIMMS
Benders Decomposition Applied to Cooperative Lot-Sizing, Dezember 2014.

395. W. GRAUBERGER, A. KIMMS
Revenue Management under Competition within Airline Alliances, November 2014.

394. W. GRAUBERGER, A. KIMMS
Airline Revenue Management Games with Simultaneous Price and Quantity Competition, Oktober 2014.

393. A. KIMMS, W. GRAUBERGER
Computing Pure Nash Equilibria in Network Revenue Management Games, September 2014.

2013

392. D. BERGENROTH, C. BREITBACH, A. DZHAMAY, J. KAUFMANN, K. MAHLER
Social Business Intelligence: Entwicklung, Architekturen und Perspektiven, Dezember 2013.

391. A. KIMMS, W. GRAUBERGER
Computing Most Preferable Pure Nash Equilibria in Network Revenue Management Games Within Alliances, November 2013.

390. A. KIMMS, I. KOZELETSKYI
Core-Based Cost Allocation in the Cooperative Traveling Salesman Problem under Rolling Horizon Planning, Juli 2013.

389. C. DOBNER, A. KLEIN
Eine kritische Analyse des Status-quo der Forschung zum Neuromarketing, Juli 2013.

388. A. KIMMS, C. REINERS
The Vehicle Routing Problem with Time Windows, Collaboration Between Freight Carriers, and a Rolling Planning Horizon, Februar 2013.

2012

387. D. ZIELING, A. MAHAYNI, S. BALDER
Performance evaluation of optimized portfolio insurance strategies, November 2012.

386. A. KLEIN, C. GÖBEL
Die Eignung von Verfahrensvarianten der Conjoint-Analyse zur Gestaltung von Produktbündeln, November 2012.

385. L. OCHTROP, A. KIAUSCH
Green IT - Eine ressourcenorientierte Sicht auf das Informationsmanagement, Oktober 2012

384. W. GRAUBERGER, A. KIMMS
Computing Inexact Nash Equilibria in General Network Revenue Management Games, Oktober 2012.

383. P. WICKER, J. PRINZ, D. WEIMAR
Big spenders in a booming sport: Consumption capital as a key driver of triathletes' sport-related expenditure, September 2012.

382. J. PRINZ, D. WEIMAR, P. WICKER, C. DEUTSCHER
No pain, no gain: Effort and productivity in professional soccer, September 2012.

381. D. WEIMAR
Facebook-Fans und Konsumkapital in der Unterhaltungsbranche, September 2012.

380. A. KIMMS, I. KOZELETSKYI
Cost Allocation in the Cooperative Traveling Salesman Problem under Rolling Horizon Planning, August 2012.

379. LI, YUAN
Downward Accountability in Response to Collective Actions: The Political Economy of Public Goods Provision in China, Juli 2012.

378. T-M. THEIS, S. NEUHAUS
Umsetzung ausgewählter Supply-Chain-Operations-Reference-Metriken durch das EPCglobal Network, Mai 2012.

377. D. STEUTEN
Modelling and Forecasting German Mortality using the Lee-Carter Mortality Model and Multi-Factor Extensions, April 2012.

376. A. ELIAS, R. KOPP, P. KOTULA, M. MAIWALD, A. KIMMS
The Capacitated Split Delivery Vehicle Routing Problem for Workforce Scheduling in a Nuclear Contaminated Environment, Januar 2012.

375. R. HEIM, M. LINDEN
Konzeption eines Rahmenwerks zur Gestaltung und Bewertung von Geschäftsmodellen, Januar 2012.

2011

374. S. KOHNE, M. LINDEN
Eignung von Verfahren der Mustererkennung im Process Mining, Oktober 2011.

373. C. VAN DER SLUIJS, M. LINDEN
Einsatzmöglichkeiten von Analyse- und Bewertungsverfahren zur Leistungsmessung im Geschäftsmodellmanagement, Oktober 2011.

372. D. STEUTEN
Deferred Variable Annuities with Guaranteed Minimum Death and Income Benefits under Stochastic Mortality and Investment Risk, September 2011.

371. A. KIMMS, K. MAASSEN
Cell-Transmission-Based Evacuation Planning with Vehicles and Pedestrians, Juni 2011.

370. D. ÇETINER, A. KIMMS
Selfish Revenue Sharing Mechanisms for Airline Alliances, Februar 2011.