Structuring Collaborative Business Intelligence: A Literature Review

Jens Kaufmann
University of Duisburg-Essen, Germany
jens.kaufmann@uni-due.de

Peter Chamoni
University of Duisburg-Essen, Germany
peter.chamoni@uni-due.de

Abstract

Cooperative work in the processes of analysis and decision support currently gains strong attention in the business world. This is motivated by spreading corporate structures and technical developments like social media and network-oriented data storage that encourage the users’ comprehension and demand for easy communication about data. This article reflects the state of research in the domain of business intelligence regarding the opening of processes for new data sources and analysts. Existing approaches are often labeled collaborative business intelligence (CBI), but differ heavily in definitions and focus. Therefore, a framework is presented and three main fields for the research of CBI are identified, which encompass internal communication, data storage with (external) partners and data analysis with partners. This article compares the findings with developments on the software market and describes open topics in the research domain.

1. Introduction

The development of Decision Support Systems (DSS) goes back to the 1960s, when IT systems were used for first computer aided decisions [53]. Since then, the understanding of DSS has broadened, and different categorizations for the system class have evolved [2, 34, 53]. Since the early 1990s, Business Intelligence (BI) is the commonly used term for analytic applications and partly for the underlying or supporting IT systems [76]. In this context, the most often used concept for data extraction, storage, and presentation is the Data Warehouse (DWH), where centrally stored (or at least accessible) data can be used by different applications for analytical or reporting purposes [35, 36].

Over the past years, building powerful BI systems has become a major concern for companies all over the world [76] and market research institutes like Gartner publish reports that particularly cover the research and development in this field, e.g. by using the hype cycle concept [7, 8]. According to these reports, an emerging, yet quite undeveloped topic of interest is Collaborative Decision Making (CDM), which includes Collaborative Business Intelligence (CBI). From Bitterer’s point of view [7], CDM is a technical solution, while its usage in literature is of a more general understanding, describing the decision process of groups, hence often considered as Group Decision Making [54]. Specialized approaches exist e.g. for cross-company planning and forecasting [28]. CBI on the other hand focuses on the joint execution of BI tasks and can be understood as the enrichment of BI systems by functions known from social software [7] – a restrictive definition we will discuss in this review (see section 4).

Social software describes a class of systems which purpose is to support communication and collaboration between different parties [21]. Among those, Online Social Networks (OSN), take an important role, as their abilities exceed the possibility to only exchange content, as e.g. wikis do [11, 37, 56]. OSN are web services that allow users to create personal profiles, to connect to others, and to track these connections over several levels [12, 48]. The use of OSN in private areas raises the users’ desire to have access to similar communication patterns in business. The analysis of social media data, e.g. from OSN, is often labeled Social Business Intelligence (SBI) [26, 51], conflicting with the definition of SBI as the use of social software functions for collaboration in BI systems [41].

CBI itself does not follow a single definition either. Some authors focus on technical aspects and internal communication patterns [6, 7, 51], others mention cross-company approaches [31, 40] and address organizational aspects [25]. The understanding of CBI and the respective field of research are disorderly. To assist further research, the current state of research should be examined and structured, as this is an accepted and valued means in general scholarship and the field of information systems [73].

This article therefore targets two main questions in the field of CBI:

1) What definitions of CBI exist and how popular are they?
2) Which aspects are primarily targeted by researchers and which areas are promising for future research?

To answer these questions, we conducted a literature review and structured the field of CBI using a straightforward framework, identifying the main understandings and topic areas.

2. Research Methodology

Literature reviews have been a widely accepted and often used means to identify and describe the state-of-the-art in a field of research [5, 19, 29, 77, 81]. We chose the methodology of Fettke [29] to structure our review and supplement it with a definition of the review scope based on Cooper [19] as a recommended means by vom Brocke et al. [73]. The review process therefore consists of five steps:

1. Problem formulation and scope definition, which have already been given in section 1.
2. Search for literature, whose methodology we explain in section 2.
3. Evaluation of literature, which is presented by the framework in section 3.
4. Analysis and interpretation, of which we outline our findings in section 4 and 5, and
5. Presentation of the results, which is given by this article itself and its conclusion.

2.1. Definition of review scope

The scope of this literature review is shown in Figure 1. According to Cooper [19], six characteristics are mentioned; the categories applying to this article are shaded in gray. Our focus lies on the use and understanding of the terms and therefore can be labeled as set on theories. We chose a neutral representation of central issues and integrated the different perspectives within a framework. In our literature search we have been as exhaustive as possible within the borders of topic-related journals and conferences. The outcomes are presented with a conceptual order, although historical aspects will be mentioned, where appropriate. This article targets specialized and general scholars as well as practitioners.

2.2. Literature Search

As mentioned before, the term CDM is widely used and therefore cannot provide the necessary focus on decision support or BI. A search for CDM with Google Scholar provides an expectedly large yield of about 18,800 results with different fields of research, among those medicine [67] and education [13].

Thus we concentrated our search on CBI and SBI. While CBI is conceptually highly related to CDM, SBI has been included as a complement to consider research approaches that look at social media as a communication enabler in BI processes.

We took a two-step approach for gathering sources and literature, beginning with a primary search in databases and journals plus articles recommended by senior scholars. Afterwards, we conducted a backward search on the reviewed articles, i.e. we followed relevant articles from the authors’ references. At this point we allowed for slight differences in notation of the search terms, if the content of the articles was clearly relevant to the topic.

In our primary search we included all journals of the AIS senior scholars’ basket of eight, which contains all major journals in the field of IS. We furthermore included all A-ranked publications according to the journal and conferences ranking of the WKWI [80]. This ranking includes all basket-of-eight publications, but gives a broader overview of the field of IS and lists specialized journals like ‘Decision Support Systems (DSS)’. The WKWI ranking is especially useful to ensure different perspectives on the international field of IS, because it complements the mostly behavioral-scientific research in international IS by the more design-science-oriented as it is common in German speaking countries. Due to those different focuses, diverse levels of abstraction in research terms are in use [70]. Given the differences between general terms (e.g. CDM) and special terms (CBI), we this way sought to ensure a comprehensive view on definitions used. To search through the journals and conferences (e.g. International Conference on Information Systems (ICIS), European Conference on Information Systems (ECIS)), we used the databases of AIS Library, Sciencedirect and Google Scholar, which also covers a fair amount of engineering research [45]. Journals that are not covered by these databases were examined through the representative publishers, including Palgrave, Springer and Taylor Francis. Whenever the database research presented results from other journals, those have been included.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Categories</th>
</tr>
</thead>
<tbody>
<tr>
<td>Focus</td>
<td>Research methods, Research outcomes, Theories, Applications</td>
</tr>
<tr>
<td>Goal</td>
<td>Integration, Criticism, Central issues</td>
</tr>
<tr>
<td>Perspective</td>
<td>Neutral representation, Espousal of position</td>
</tr>
<tr>
<td>Coverage</td>
<td>Exhaustive, Exhaustive &amp; Selective, Representative, Central or pivotal</td>
</tr>
<tr>
<td>Organization</td>
<td>Historical, Conceptual, Methodological</td>
</tr>
<tr>
<td>Audience</td>
<td>Specialized scholars, General scholars, Practitioners, General public</td>
</tr>
</tbody>
</table>

Figure 1. Literature review taxonomy based on Cooper [19]
Because we aim for our work to point out the understanding and usage of the relevant terms, we acknowledge that the meaning of a word is defined by its usage in the language [79] and have extended our search to articles of popular science. We reviewed the first one hundred results of a Google web search, but restrained ourselves to editorially supervised pages. Hence, vendor publications were not used, as they are presumed to have a marketing focus. A separate view on vendors and market analysts will be given in section 5 of this article.

In any case, we searched for the word order rather than for single terms and used a full-text search over all available years whenever possible. Our search was conducted and several times renewed from October 2012 to March 2013.

2.3. Overview of results

The database search with the given pattern led to 135 hits in total, showing CBI as the more frequent term (see Table 1). Of these hits, 21 were selected to actually be reviewed in detail and listed in this article. The quota of 15.6% results from four factors. First, due to the use of different databases there were double-hits, e.g. because Google scholar in partial lists AIS articles. Second, some papers in Asian languages like Chinese have not been considered if only a translated abstract was available, that did not point out the information needed for our review. Third, different hits that all belonged to a monograph, were summarized. Fourth, articles, that only mentioned CBI in a footnote without explanation, e.g. Cheng et al. [17], were left out.

An even smaller quota can be found with the results of the web search. Reasons for this are the restriction to editorially managed web sites and the removal of multiple web sites under the same domain. It is to be noted, that the reported amount of indexed websites was ca. 21,100 results for CBI, and about 4.79 million results for SBI. Because not all articles could be reviewed, we can only assume that the term ‘social’ had severe influence on this, probably promoted by the general interest in social media.

Based on the results in databases, the number of articles was raised as described before, leading to 40 publications in total. It is noticeable that the number of publications rose strongly in 2010 and the following years. An increasing interest thus can be safely assumed, which could be strengthened by a general advance of social media or BI research.

3. Framework for classifications

As stated before, current definitions of CBI differ notably. We propose a framework that aids understanding followed by a classification proposal. To prepare a research agenda all publications are additionally considered under different aspects (or perspectives).

3.1. Dimensions of the framework

The framework consists of two dimensions that are derived from the general understanding of BI which covers the process from data extraction to information access for analysis [53]. According to this differentiation, both data provision and data analysis can be broadened by collaboration. This is done by adding new sources that normally are not taken into account in BI systems or by adding new analysis partners, e.g. other departments or companies. Schlagwein et al. [63] distinguish the openness of information sources by extending the access rights from exclusive access via group-oriented access to open access. If that model is applied to a BI system, this could include new sources (taking data) or new analysts (giving data).

The first dimension accordingly describes the opening of a ‘classical’ BI system to new sources, beginning with no change. In this stage, only the usual source systems like enterprise resource planning systems or production planning systems are used. An initial opening occurs, if company-internal data is added that is not explicitly available in existing source systems. Examples are employees’ knowledge for prognosis [66] or context-sensitive information that gets more often used as mobile decision support (‘mobile BI’) evolves [16]. If trust between different companies is high enough [40], a combination of data sources becomes an opportunity. In difference to adding external data to the companies’ own databases,
strategic alliances or governmental organizations share data via networks [44, 58]. The strongest extension of the data provision process occurs if completely external and independent sources are brought in. Concepts to encourage data reception from private or public organizations can be found under the term ‘open data’ [1]. Alternatively, indirect sources can be taken into account, e.g. the access to data of OSN [50, 61].

The second dimension refers to the participating analysts in the BI process. Usually, these are employees of one or more departments, e.g. purchasing, sales, controlling, or a specialized BI competency center [47]. If it is of need to (partly) include specialists of other departments, the BI process opens up within the enterprise. Currently available BI systems therefore use communications instruments that are coupled with reports. That allows for linkage of reports and integrated communication over figures [22, 72]. Corporate networks and alliances can use shared analysis processes under the assumption that valuable information is only brought up if the amount of data is bigger than single companies can obtain. Insurance companies, e.g., can exchange data to fast gather information about natural catastrophes or fraud patterns [40]. Fully opening access to the analysis process requires the inclusion of third-party workers.

Böhringer et al. [9] discuss this scenario and refer to crowdsourcing, i.e. handing out single tasks to an anonymous crowd via the internet. Admittedly, at this moment, neither ‘BI Crowdsourcing’ nor ‘Crowd BI’ have great impact in the research field.

3.2. Classification of publications

Previous thoughts lead to the actual two-dimensional framework that presents the opening of data provisioning on the vertical and the opening of data analysis on the horizontal axis. We observed that some publications contained different definitions and therefore cited them twice. Figure 3 shows 40 publications with 43 points due to the multiple definitions. Only those publications are shown that explicitly used the search terms and/or provided a clear definition. Renewed definitions of the same author were allowed (e.g. [7, 8]). The framework also shows the initial search terms that led to the findings. All articles that refer to CBI or CDM are shown within a rectangle; all articles referring to SBI, Social Analytics, Social Media Analytics or Socialytics are shown in ellipses. If several articles show the same understanding of the term, they are put together in one bigger rectangle or ellipsis.

![Figure 3. Framework for literature review of CBI and SBI](image-url)
Acknowledging the initially descriptive character of the framework, two bigger clusters of publications can be seen (external data provision with internal analysis; existing data use with company-wide and partly partner-oriented analysis). They show a rough, widely-accepted distinction between SBI and CBI. Especially concerning partner-oriented collaboration, different articles are named in the context of CBI, as we describe in section 4.

3.3. Social Business Intelligence

A literature review with the sole focus on SBI has been conducted by Dinter and Lorenz [26], who state that there is no unambiguous understanding of the term and follow the definition of Zeng et al. [82] for the term ‘social media intelligence’. They transfer it to SBI and state it “(…) aims to derive actionable information from social media in context rich application settings, develop corresponding decision-making or decision-aiding frameworks, and provide architectural designs and solution frameworks for existing and new applications (…)”[26].

Our classification of publications into the framework supports this understanding of SBI and distinguishes it quite clearly from CBI. The most comprehensive definitions cover the analysis of data of all available social media in context of decision support [43, 51, 61], sometimes referring to real-time analysis [33]. Wailgum [74] focuses on OSN, supported by Roe [59], who states that users now expect the kind of personalized advertising made possible by these analyses. Empirical work focuses on the short message service Twitter [10, 50]. Technical approaches focus on data modeling and system architectures [20, 60], while organizational aspects are sparsely covered and describe forms of project staffing [30] or the use of employees’ knowledge [66].

4. Collaborative Business Intelligence

As it can be seen in the framework, the understanding of CBI is heterogeneous. Hence a more differentiated examination has to be given to the general statement: Collaborative Business Intelligence describes an extension on purpose of the analysis process in BI by adding internal and external partners and necessarily opening access to data. Only two publications differ from this view and focus on the general process of information gathering [38] and the use of specialized software agents [49].

Based on the framework we initially decide between three basic trends in the general understanding. Those are internal communication (IC), where mostly existing systems are used, partnership in data (PD), where external partners get involved in the data provision process, and partnership in analysis (PA), where partners work together in the data analysis process.

To sort the publications, a matrix is used, comparing the classification by the framework to the following aspects: technical, organizational, functional, and economical (see Table 2). We hereby follow a scheme that is applied in the field of BI (concerning TFO) [9, 15] and is very similar to the one suggested by Turban for structuring the research fields in CBI [72]. The latter one complements the economical aspect into the scheme. Table 2 presents the number of publications that focus on each field in the matrix. Articles that deal with definitions rather than extensive research are sorted in to a field in respect to their particular derivation of the definition.

![Table 2](image)

4.1. Internal Communication (CBI-IC)

The majority of publications understand CBI as a primarily technical-driven development. Articles published before 2008 refer to a situation without the massive use of OSN (see Richter et al. [56] for information about OSN usage). First publications describe an extension of BI portals and workflow components with possibilities for real-time communication and joint data storage [18, 55, 68, 69]. All this work focuses on the presentation layer of BI systems and authors who understand CBI as a communication enabler act equivalently. They emphasize the possibilities for annotations of figures in reports, direct feedback and consultation, and the simple user interfaces to combine different knowledge bases [7, 27, 39, 42, 43, 51, 52, 71, 72]. An exemplary definition can be found in [51]: “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 (…)”. Devlin [25] considers organizational changes, especially how decision processes can be improved by comparing the process and the outcome.

Two publications exist that mostly target the technical aspect, but also look at the other aspects.
Dayal et al. [22] use the example of a data center to show how different experts can collaborate by supervising performance indicators. To enable this, the standard DWH scheme is broken up and all data is not – as is usually done – integrated into one database, but instead can be accessed via a new layer (“ontology and meta data layer”). The authors discuss organizational and functional aspects, especially the challenge to create good ontologies to connect all different data sources and make them analyzable. They review a practical case and name economic reasons for their CBI platform time- and cost-saving potential by coupling operative real-time systems to BI systems. Berthold et al. [6] do not refer to the classical DWH architecture, but present a model, that can be transferred to it. In their model, data is integrated by a separate layer to another layer for collaborative (ad hoc) analysis. They decide between “infospaces” for the (conceptual) scheme of data and “dataspaces” for the actual storage of the data. From an economical perspective, self-service BI is mentioned as the main driver for collaborative analysis, e.g. the evaluation of reports to generate recommendations for the users.

An empirical study of Liu et al. [41] tries to evaluate the cohesion between CBI and success in decision making. They discover a slightly positive effect, but remark that most of the study participants where Chinese students.

4.2. Partnership in data (CBI-PD)

The simplest form of extending the BI process over the boundaries of the own enterprise is to gather external data. Thus Schwalm and Bange [65] describe CBI as a technical topic that deals with the structured exchange of XML-based data. Alwis [3] describes CBI as the business goal to integrate external data.

A clearly more partnership-oriented approach that considers the collective storage of data is described with the business intelligence networks (BIN) by Golfarelli and Rizzi [31, 58]. The economic foundation is – from their point of view – the need for detailed knowledge about the enterprises’ environment. The article shows technical variants for shared data storage and distinguishes three alternatives for ad hoc-analysis (OLAP): First, an integrated DWH; second, a federative approach with a master data scheme, valid in all DWH; third, a peer-to-peer approach - for which functional challenges in converting different schemes are extensively discussed.

Another federative approach is given by Martins et al. [44]. Again, a central ontology is used to semantically connect the different DWH schemes. They describe different “environments”, “layers”, and “cycles” to create a technical-oriented reference architecture. The actual development of this ontology, however, is explicitly not the topic of their article. They present the joint usage of data by Brazilian government institutions as an exemplary practical case for this form of CBI.

4.3. Partnership in analysis (CBI-PA)

Few publications address analyses based on partnership, i.e. across company borders. A technical architecture with a central DWH is described by Mettler et al. [46]. Access to this central DWH and synchronization with local DWH is controlled by security protocols of already shared systems for joint purchasing processes. The aim is to keep usually fluent data from different operative processes and to analyze it to get better insights to the process as a whole. A similar idea is presented with the value chain cockpit by Werner et al. [78]. The authors describe the (technical) integration of systems for communication with customers and suppliers. BI systems are mentioned on the presentation layer, as analyses can be conducted along the whole value chain. Different from this vertical integration, Liu and Daniels [40] discuss joint analyses of multiple companies in the same area of business. Their exemplary application for risk assessment is the system FISH (Fraud Information System Holland), used by insurance companies in the Netherlands. A technical solution is not provided; functional integration of the data is assumed.

5. Perception of and Agenda for CBI

In addition to a literature review on research publications, we take a brief look at CBI as a topic for vendors, market analysts and practitioners. To answer the question for a future research agenda, we evaluate the probable sustainability of the topic and discuss a possible agenda based on our findings.

5.1. Perception by vendors, market analysts, and practitioners

Data analysis processes tend to decentralize in companies, involving more and more different users [32, 61, 64]. This development raises the desire for easy collaboration in BI environments. Established vendors provide CBI functions in their systems that correspond to the pattern of internal communication, but do not support cross-company analysis [62]. Conversely, a Gartner study shows that only ten percent of social media initiatives that aim to foster communication in companies are successful [57] which motivates future work concerning economic questions.
Although we did not conduct a full complementary survey amongst practitioners, we verified their interest in CBI by single interviews with senior BI consultants and managers who are trusted to be vendor-neutral. All of them stated that bi-directional communication over data and analyses is the main focus of their understanding of CBI. They also added that cross-company analyses are not a current practice, but will probably become more valuable in terms of joint interpretation of analysis results within the next years. That concurs with our presented review results. Anticipated to our suggestion for a research agenda, we assume that a profound analysis of practitioners’ views could bring valuable insights about real usage and future demands of CBI – especially, when compared to the market analysts’ point of view as shown above.

5.2. Sustainability of interest in the field of collaborative decision making and CBI

As stated in the introduction, BI, and IS in general, are vivid fields of research and practice, which often deal with new and changing topics. We therefore examined, whether the area of collaborative work in decision making processes and analyses is of sustainable interest or only a passing fad [70]. As CBI is a promising and clearly quite young topic, we chose different terms and compared their use in research of the recent past. The terms chosen were CDM and Group Decision Making (as they are more general terms, see section 1 of this article) as well as Group Decision Support Systems and Group Support Systems [75], which purpose is to guide through a structured collaborative decision process [24]. We analyzed publication databases we had already used for the literature review, as they cover most of the high-ranked journals and conferences (e.g. AIS: ICIS, ECIS, MISQ; Sciedirect: Information Systems, Int. Journal of Inf. Mgmt.). As it is shown in Figure 4, AIS shows a steady number of publications per year over recent years regarding the given topics. Sciedirect, the database with the far most hits, shows continuously growing publication numbers, which can be taken as a very positive sign for interest in the field.

5.3. Research agenda

We identified three different forms of CBI (IC, PD, PA) and examined all of them by four aspects or perspectives (TOFE). To assist others in finding most promising spots for research and improvement, we use the aspects’ view to identify those spots, as we believe that the most pressing challenges are quite alike for all three forms.

(a) Economically seen, all forms of CBI are under pressure to justify their usefulness. Vendors’ ambitions are in question due to recent studies of little impact on communication patterns [57]. Other empirical studies hardly exist [41]. Partnership-oriented approaches mostly refer unclearly to the expected benefits from sharing data along the value chain. It should be examined, whether existing and proven methods like Collaborative Planning, Forecasting and Replenishment (CPFR)[28] can be applied to BI.

(b) Most publications focus on technical aspects, but only a few exceed the implementation of social-media-like functions into BI systems [6, 22, 58]. Current research must explore, how different layers of BI systems will undergo change in CBI environments. At the level of data provision and storage, central and decentralized approaches should be compared in detail. Their integration in existing networks must be reviewed under the premise of keeping low access times. This is of special concern with the upcoming challenges of ‘big data’, i.e. the processing of large amounts of fast-changing, poly-structured data [14]. Also missing is a thoughtful comparison of all available functions in social media and their usefulness for CBI systems.

(c) Functional aspects that should ensure meaningful and clear interpretation of analyses or query results on the data are largely ignored in current research. Approaches for internal co-work simply assume the consistency of data, but even in single companies, this has to be created by an integrated DWH first. When systems of different companies are coupled, the problems evolve. Ontologies are mentioned, but their creation is not in focus [22, 44]. When different schemes are matched, the risk of
information loss is present. There have to be measures defined that evaluate how much information can be preserved under these circumstances [58].

(d) At last, organizational aspects have to be reviewed. It seems reasonable to transfer insights in group decision processes to recent BI systems. Impacts of cross-company systems may especially target the concept of a BI competency center [47]. Finally, authority over data and security (physically and concerning access rights) has to be discussed in detail.

6. Conclusions, limitations, and outlook

Motivated by the latest developments in the field of BI, this article gives a comprehensive and structured overview about the understanding and usage of CBI in literature. In this context, similarities and differences to the topic of SBI are pointed out.

To answer the question, which definitions exist, we identified three main understandings of CBI, based on a widespread literature review and a straightforward framework: CBI-IC is the enrichment of existing BI systems by communication tools as known from social software. CBI-PD extends data provision across company borders to shared data platforms. CBI-PA defines CBI as joint data analysis, conducted by different companies. The results show a clear focus on the aspect of internal communication and very few ideas on how to handle or to use cross-company analysis. We stated that in every aspect (technical, organizational, functional, and economic) current approaches cannot fully cover all open questions, but technical progress seems to be the most energetic driver for the field.

An obvious limitation of our work is the number of articles reviewed. Although we thoroughly gathered a high number of articles and publications, the review was based on only a few search terms. That is due to the huge amount of results that return when looking into less specific terms such as CDM. An exhaustive literature review would have to set many of those vague terms in relation to decision support or BI. Besides, a more detailed review of current vendor activities could lead to more insights about the most likely understanding of the terms in the near future or of economic justifications to come. This should be accompanied by a review of non-published views on the field of CBI, e.g. by interviews or case studies with practitioners.

Considering all developments and remarking that the general interest in this field of research is stable or even rising, we assume that future research in the field of CBI will fall on fruitful ground.

7. References
