Advances in Business-Related Scientific Research Conference 2015 in Milan
(ABSRC 2015 Milan)
December 10-11, 2015, Milan, Italy

BIG DATA - A NEW VALUE FOR ORGANIZATIONS

Celina M. Olszak*
Katowice University of Economics
Poland
celina.olszak@ue.katowice.pl

Maria Mach-Król
Katowice University of Economics
Poland
maria.mach-krol@ue.katowice.pl

Abstract

The research objective for this study is to investigate some values and benefits that Big Data may offer for an organization. Firstly, the issue of Big Data is identified. Then, Big Data application areas are presented. Finally, the most important techniques for analyzing Big Data are discussed. The study was based mainly on a critical analysis of literature, creative thinking, and an interpretive philosophy. The results of this research can be used by IT, business leaders, and policy makers as they plan and develop Big Data applications in their organizations. The paper provides valuable information on the chances and the possibilities of Big Data applying in organizations.

Keywords: Big Data, Business Intelligence, analyzing Big Data, organization, value

Topic Groups: Organizational information and communication systems, Technology and innovation management

INTRODUCTION

Organizations collect and analyze information for different purposes. They may refer to (Ishikawa & Nakagawa, 2013): (1) deepening and expanding of existing business, (2) risk management, (3) information control and security control and management, as well as (4) information manipulation. Until now, organizations usually have gathered and analyzed internal information in order to better understand their business processes and to improve decision-making on operational and tactical level (Negash & Gray, 2008). Recently, they have started to collect and explore the external information that comes from the organizational environment, e.g.: from competitors, industries, government, administration, and healthcare (Baaras & Kemper, 2008; Olszak, 2013). Such data called "*Big Data"* (BD) refers to databases whose size is beyond the ability of typical databases software tools to capture, store, manage, and analyze. The researchers find that processing and analyzing such data can create

significant value for organizations and for the world economy, enhancing the productivity and competitiveness (Manyika et al., 2011). Many authors emphasize that organizations that are able analyze BD will win in the rapidly globalizing information society (Davenport & Harris, 2007; Ishikawa & Nakagawa, 2013). They predict that BD will revolutionize management tools and will be important for the success of any organization (Weiss, 2002).

The main aim of this study is to investigate some values and benefits that BD may offer for an organization. The paper is organized as follows. Firstly, the issue of BD is identified. Then, BD application areas are presented. Finally, the most important techniques for analyzing BD are discussed. The study was based mainly on a critical analysis of literature, creative thinking, and an interpretive philosophy. The results of this research can be used by IT, business leaders and policy makers as they plan and develop BD applications in their organizations.

RELATED WORKS

Big Data issue

The interest in analytics in information systems domain has continued for many years. This is reflected in research that has been conducted in the 80's and 90' of XX century and that has referred mainly to Management Information Systems, Decision Support Systems, Expert Systems, and Executive Information Systems (Liautaud & Hammond 2002; Moss, & Atre, 2003; Davenport & Harris, 2007; Negash & Gray, 2008; Wixom & Watson, 2010). XXI century is characterized by the development of data warehousing, On Line Analytical Processing (OLAP), Business Intelligence (BI), Competitive Intelligence (CI) and last time Big Data (BD). The common assumption underlying these analytical systems is that an appropriately acquired, collected, analyzed, integrated, and used information may be a critical component for the success of the organization. Analytical systems may help organizations to reach strategic goals, make better decisions, improve business processes, increase profitability, and improve customer satisfaction (Weiss, 2002; Williams & Williams, 2007; Howson, 2008; Olszak, 2014). At the same time it is emphasized that this challenge becomes more difficult with the constantly increasing volume of information, both internal and external to an enterprise.

To better understand the idea of BD is worth recalling the basic assumptions underlying the idea of BI and CI. They present last stages in the evolution of analytical systems and are strong associated with BD. It is stated that BI is focused mainly on gathering and analyzing internal information, thus it helps to improve internal business processes and decision-making on an operational and tactical level of organizations (Negash & Gray, 2008). CI refers to collection and exploration of external information that comes from the organizational environment (Baaras & Kemper, 2008; Karim, 2011; Olszak, 2014). This includes information about competitors, industries, governments and nations for management strategy, as well as on changing business environment and trends of suppliers of materials, exporting nations, competitors and overseas markets.

Steyl (2012) argues that BI cannot replace CI and vice versa. BI is the management of a company's internal data and processes that are well structured. This includes the Extract, Transform and Loading (ETL) of this data in a well-designed data warehouse in order to extract and present the data as intelligent information. BI may help any business make better

decisions based on its own historic information and can therefore project possible future trends. Typical BI tools include: reporting, OLAP, data warehouse, data mining and visualization tools. In turn, the data that originates from external sources, distinctive for CI, is usually semi-structured and unstructured.

According to Ferguson (2012) the term "*Big Data"* is "associated with the new types of workloads and underlying technologies needed to solve business problems that we could not previously support due to technology limitations, prohibitive cost or both". BD is therefore not just about data volumes but about analytical workloads that are associated with some combination of data volume, data velocity and data variety that may include complex analytics and complex data types. BD refers (Manyika et al., 2011) to datasets whose size is beyond the ability of typical database software tools to capture, store, manage, and analyze.
BD, compared to BI and CI, concerns even broader spectrum of data and more complex sophisticated data processing (Table 1). It generally refers to data that exceeds the typical storage, processing, and computing capacity of conventional databases and data analysis techniques (Manyika et al., 2011). BD requires tools and methods that can be applied to analyze and extract patterns from large-scale data. The rise of BD has been caused by increased data storage capabilities, increased computational processing power, and availability of increased volumes of data, which give organizations more data than they have computing resources and technologies to process. BD concerns mainly unstructured information about competitors, customers, public sector, transport, healthcare and other stakeholders of the organization.

Table 1: The basic differences between BI, CI and BD

	Business Intelligence	**Competitive Intelligence**	**Big Data**
Purpose	Analyzing mostly of the internal business processes, improvement operational and tactical decisions	Monitoring, analyzing of external environment – mainly competitors	Analyzing of the whole environment of the organization: internal resources, customers, suppliers, users of Internet, communities of practices,
Scope	Organization	Environment of organization, mainly competitors	Whole environment of the organization
Content/data	Well structured content, internal data coming from databases, ERP, transaction systems	Semistructured, unstructured content, external data coming from public, open resources, Internet	Unstructured content, external data coming from public, open resources, Internet, mobile devices, social media
Used tools, technologies	OLAP, data mining, data warehouse	Advanced data mining, predictive modeling, web mining, text mining, exponential random graph models, search based applications, dashboards, SOA	Advanced data mining, predictive modeling, web mining, opinion mining, text mining, exponential random graph models, search based application, dashboards, SOA, Hadoop, MapReduce, parallel processing, real-time processing, machine learning techniques

RESEARCH METHODS

The main objective of this paper is to identify the possibilities of BD usage and to explore values that BD may offer for organizations. The different research methods and techniques are used to achieve the specific objectives and research tasks. They include in particular: a method of critical analysis of literature, as well as creative thinking, and an interpretive philosophy. The research is carried out on the basis of secondary and primary sources of information. Secondary sources include literature from the field of strategic management, computer science, and management information systems. In addition, research reports and expertise opinions from different research institutions are used. The primary sources base mainly on own experiences and experiments with the work of different techniques, and software programs for BD analyzing.

FINDINGS AND DISCUSSION

Big data application areas

It is believed that BD analyzed in combination with traditional enterprise data (most structured and semi-structured), enables organizations not only to better understand their business, but first of all to change it and to have new sources of revenues, more stronger competitive position and greater innovation.

Nelson (2010) argues that the implementation of BD in management will contribute mainly to:

- decisions, facts and context will be developed through crowdsourcing;
- data and reports will incorporate narrative context information supplied by users;
- data will have a more direct linkage with action;
- people will be able to directly act on information;
- business decisions shall be monitored so that interventions and our hypotheses about business tactics will be tagged in the context of the data that measures its effect;
- visualization data and complex relationships will be easier and more intuitive models of info-graphics will become mainstream;
- the ability to detect complex patterns in data through automated analytic routines or intelligent helper models will be built into analytic applications;
- finding information will be easier and search results will provide context so that we know when we have the right results.

According to McKinsey Global Institute (Manyika et al., 2011) BD creates value for an organization in several ways. They include mainly:

- creating transparency – making relevant data more readily accessible across otherwise separated departments can sharply reduce search and processing time;
- enabling experimentation to discover needs, expose variability, and improve performance – organizations can collect more accurate and detailed performance data on everything from product inventories to personal sick days;
- segmentation populations to customize actions – big data allows organizations to create highly specific segmentation and to tailor products and services precisely to meet those needs;
- replacing/supporting human decision making with automated algorithms - advanced analytics can substantially improve decision making, minimize risks and unearth

valuable insights that would otherwise remain hidden. Organizations are already making better decisions by analyzing entire datasets from customers, employees, or even sensors embedded in products.

- innovation new business models, products, and services – big data enables organizations to create new products and services, enhance existing ones, and invent entirely new business models.

BD beneficiaries may include a wide group of users starting from specialists in controlling, financial reporting, through salespeople, specialists from administration, security, up to members of the board. Sectors that may use BD most frequently include trading companies, insurance companies, banks and a financial sector, telecommunications, manufacturing companies, healthcare, and public sector (Table 2).

Table 2: Big data application areas

BD applications	Objectives
Retail industry	• Forecasting. Using scanning data to forecast demand and based on the forecast, to define inventory requirements more accurately. • Ordering and replenishment. Using information to make faster decisions about items to order and to determine optimum quantities. • Marketing. Providing analyses of customer transactions (what is selling, who is buying). • Merchandising. Defining the right merchandise for the market at any point in time, planning store level, refine inventory. • Distribution and logistics. Helping distribution centers manage increased volumes. Can use advance shipment information to schedule and consolidate inbound and outbound freight. • Transportation management. Developing optimal load consolidation plans and routing schedules. • Inventory planning. Helping identify the inventory needed level, ensure a given grade of service.
Insurance	• Claims and premium analysis. The ability to analyze detailed claims and premium history by product, policy, claim type, and other specifics. • Customer analysis. Analyze client needs and product usage patterns, develop marketing programs on client characteristics, conduct risk analysis, improving client service. • Risk analysis. Identify high-risk market segments and opportunities in specific segments, relate market segments, reduce frequency of claims.
Banking, finance and securities	• Customer profitability analysis. Determinate the overall profitability of individual customer, current and long term, provide the basis for high-profit sales and relationship banking, maximize sales to high-value customers, reduce costs to low-value customers, provide the means to maximize profitability of new products. • Credit management. Establish patterns of credit problem progression by customers class and type, warn customers to avoid credit problems, to manage credit limits, evaluate of the bank's credit portfolio, reduce credit losses. • Branch sales. Improve customer service and account selling, facilitate cross

	selling, improve customer support, strengthen customer loyalty.
Telecommunications	• Customer profiling and segmentation. Determine high-profit product profiles and customer segments, provide detailed, integrated customer profiles, develop of individualized frequent-caller programs, determine future customer needs. • Customer demand forecasting. Forecast future product needs or service activity, provide basis for churn analysis and control for improving customer retention.
Manufacturing industry	• Sales. Provide analyses of customer-specific transaction data. • Forecasting. Forecast demand, define inventory requirements. • Ordering and replenishment. Order optimum quantities of items. • Purchasing. Helping distribution centers manage increased volumes. • Distribution and logistics. Can use advance shipment information to schedule and consolidate inbound and outbound freight. • Transportation management. Developing optimal load consolidation plans and routing schedules. • Inventory planning. Identify the inventory level needed, ensure a given grade of service.
Healthcare	• Consolidation of clinical, financial and operational information. BD enables to integrate and analyze clinical, administrative and financial data, which also serves to increase the efficiency in the data/work flow. • Efficiency improvement. Users can access any type of information with a fast and consistent response time, independent of the data volumes analyzed or questions asked. • Improved patient treatment and care. By means of BD, healthcare professionals have easy access to patient's data and they can create a variety of classifications/reports based on demographic data, sex, age and so on. Thanks to the evidence based medicine and capture of medical history of the patient, doctors can accurately diagnose and apply efficient treatment with reduction of risks during treatment. • Reduction of medical errors and improved patients safety. BD applications can support a larger healthcare system by the exchange of medical information on a patient. • Improved monitoring. Monitoring of the consumption of drugs, medical supplies, use of medical equipment, medical personnel, movement of patients.
Public sector	• Segmentation populations to customize actions. Tax agencies may use BD to segment individual and business taxpayers, separating them into categories and classes for examination and collection activities. • Reduction in fraud and error. • Recruiting and training talented personnel. • Operational efficiency savings. • Higher quality services. Citizens and business can spend less time and effort in their interactions with government agencies and receive services better targeted to their needs. • Increase public sector accountability, a better-informed citizenry.

Techniques for analyzing big data

Different types of techniques may be distinguished for analyzing BD. They refer mainly to: predictive modeling and advanced data mining actives, text and web mining, agent-based models, exponential random graph models, search based application, genetic algorithms, neural networks, network analysis, spatial analysis, and visualization tools. These techniques have a large impact on the functionality and possibilities of BD applications and consequently on the quality of decision making, personalization of products and services and improving business processes, and customer relationship management (Chen et al., 2012; Olszak, 2014).

Predictive modeling and advanced data mining activities constitute an interactive process aimed at the analysis of large databases, with the purpose of extracting information and knowledge that may prove accurate and potentially useful for knowledge workers engaged in decision making and problem solving (Sauter, 2010; Vercellis, 2009). Data mining activities can be subdivided into two major investigation streams, according to the main purpose of the analysis: interpretation and prediction (Poul, Gautman & Balint, 2003). The purpose of interpretation is to identify regular patterns in the data and to express them through rules and criteria that can be easily understood by experts in the application domain. The purpose of prediction is to anticipate the value that a random variable will assume in the future or to estimate the likelihood of the future events. Different mathematical models and methods are used in such advanced data mining. The most important include: regression, time series, classification, association rules, and clustering (Larose, 2005; Tan, Steinbach, & Kumar, 2005; Han, Kamber & Pei, 2011).

Text and web mining are commonly used to process and analyze unstructured web contents. Text analytics is required and used because a significant portion of the unstructured content collected by an organization is in the textual format, from e-mail communication and corporate documents to web pages and social media content (Chen, Chiang & Storey, 2012). Text analytics is based on the knowledge originating from information retrieval and computational linguistics. In information retrieval, document representation and query processing are the foundations for developing the vector-space model, Boolean retrieval model, and the probabilistic retrieval model. While in computational linguistics, statistical natural language processing (NLP) techniques for lexical acquisition, word sense disambiguation, part-of –part tagging (POST) and probabilistic context-free grammars have become important for representing text. Text mining may refer to: (1) Data processing: preprocessing of the data to the needed format; (2) Concept extraction: extraction of important concepts and terms through initial text analysis; (3) Narrative analysis: writing a narrative analysis to identify patterns and co-occurrences of identified concepts; (4) Automatic categorization: developing an automated solution; and (5) Ontology building: building ontology for future CI analysis.

Leveraging the power of big data and statistical NLP, text analytics techniques have been actively pursued in several emerging areas, including information extraction, topic models, question-answering (Q/A), event detection (ED) and opinion mining (OM). Information extraction aims to automatically extract specific kinds of structured information from documents. As a building block of information extraction, NER (named entity recognition) is a process that identifies atomic elements in text and classifies them into predefined categories. NER techniques have been successfully developed for news analysis and biomedical applications (Witten et al., 2011).

Question answering (Q/A) systems rely on techniques from NLP, informational retrieval, and human-computer interaction. Primarily designed to answer factual questions (who, what, when, and where kinds of questions), Q/A systems involve different techniques for question analysis, source retrieval, answer extraction, and answer presentation (Maybury, 2004). Event detection focuses on identifying information about events, such as type, time, place, participants and date of the event. Examples of a business event appearing in a newspaper could be a company establishing a new production facility or releasing a new product. Event detection process consists of three steps: event topic reasoning, event property extracting, and similarity comparison. The event topic reasoning step includes: representing the input text by a feature set pertaining to each event topic (e.g., when, who, where, what) and classifying the text into an appropriate event topic based on the event categorization patterns. Event property extraction refers to creating by extracting the event properties (e.g. participating company's names, dates, time, place of the event) based on event ontology. Event similarity comparison refers to the process of determining whether a new input document discusses a new or a previously known event. Opinion mining, also known as sentiment analysis (SA), refers to the computational techniques for extracting, classifying, understanding, and assessing the opinions expressed in various on line news sources, social media comments, and other user-generated content (Chen, Chiang & Storey, 2012; Newman, 2010). It attempts to automatically measure human opinions from a text written in natural language.

Web mining is used to process and analyze unstructured web contents, based on XML, Internet protocol (HTTP, SMTP) APIs (application programming interface). It enables developers to integrate diverse content from different web-enabled systems. A major component in web analytics is cloud computing, which includes applications, systems software, and hardware delivered as services over the Internet. Based on service-oriented architecture (SOA), server virtualization, and utility computing, cloud computing can be offered as software as a service (SaaS), infrastructure as a service (IaaS), or platform as a service (PaaS). The SaaS model means provision of certain software functionality. In model IaaS the infrastructure is delivered to customer. The PaaS model refers to an infrastructure and development environment for developing specific cloud applications. It is reported that there are four methods of providing cloud services: community cloud, public cloud, private cloud, and hybrid cloud (Veber, 2012).

Agent-based models. The agent is an entity that performs some actions in a particular environment and is aware of the emerging changes. Moreover, it can react to such changes (Poole & Mackworth, 2010; Olszak & Bartuś, 2013). The agent has a set of goals, certain capabilities to perform actions, and some knowledge (or beliefs) about its environment. The agent model as "a solution-oriented ensemble of capabilities including natural language processing, autonomous reasoning, proactive computing, discourse modeling, knowledge representation, action-oriented semantics, multimodal interaction, environmental awareness, self awareness, and distributed architectures" (Thomsen, 2002). It involves software entities that carry out some set of operations on behalf of a user or another program, with some degree of independence or autonomy, and in so doing, employ some knowledge or representation of user's goals and desires.

Exponential random graph models (ERGMs) are a family of statistical models for analyzing data about social and other networks (Robins et al., 2007). To support statistical inference on the process influencing the formation of network structure, ERGMs consider the set of all possible alternative networks weighted on their similarity to observed networks.

Such tools like UCINet and Pajek have been developed and are widely used for large-scale network analysis and visualization (Borgatti et al., 2002; Chen, Chiang & Storey, 2012). In addition, lightweight programming models of the current web services (HTML, XML, CSS, Ajax, Flash, J2J) and the maturing mobile development platforms such as Android and IOS have contributed to the rapid development of mobile web services (HTML5, Mobile Ajax, Mobile Flash, J2ME) in various social networks.

Search based application is a category of application that enables users to find information from any source and in any format with a search-based application. The user can enter anything he knows about the customer or product, and the search engine will try to find those customers or products that resemble the keywords entered by the user. According to Capgemini (2013), there are different reasons for using search engines, for example: (1) they are very flexible in handling any format and type of information, be it structured, unstructured or external; (2) they are able to cope with continuously evolving data structures; (3) search engines enable content-driven dimensional navigation; (4) search engines, unlike solutions based on RDBMS, are able to analyze data without the need to know the various data types; (5) end users are now quite familiar with the "Google" interface. As a consequence, they are much more independent from IT departments if they can access decision support data through a search engine; (6) when external and unstructured data is needed to support decision making, traditional data warehouse architectures are limited and search engines can help to fill the gap.

Genetic algorithms are a technique where potential solutions are encoded as "chromosomes" that can combine and mutate. This technique is used for optimization that is inspired by the process of natural evolution or "survival of the fittest" (Manyika et al., 2011).

Neural networks are computational models, inspired by the structure and workings of biological neural networks that find patterns in data. Neural networks are well-suited for finding nonlinear patterns. They can be used for pattern recognition and optimization. Examples of applications include identifying high-value customers that are at risk of leaving a particular company and identifying fraudulent insure claims (Manyika et al., 2011).

Network analysis represent a set of techniques used to characterize relationship among discrete nodes in a graph or a network. In social network analysis, connections between individuals in a community or organization are analyzed. Examples of applications include identifying key opinion leaders to target for marketing, and identifying bottlenecks in enterprise information flows (Manyika et al., 2011).

Spatial analysis is a set of techniques, some of them originate from statistics. They analyze the topological, geometric or geographic properties encoded in a data set. Often a data for spatial analysis come from geographic information systems that capture data including location information (Manyika et al., 2011).

Interactive visualization tools enable decision makers to dynamically explore ideas, investigate patterns, uncover hidden facts and share those insights across the enterprise for better decision making. An extensive suite of customizable graphical presentation options present information and insights not easily detected in tabular formats (Chung, Chen, & Nunamaker, 2005).

It is worth mentioning that there is a growing number of technologies used to aggregate, manipulate, manage and analyze big data. They include e.g.:

- Big Table - proprietary distributed database system built on the Google File System;
- Cassandra - an open source database management system designed to handle huge amounts of data on a distributed system;
- Google File System - proprietary distributed file system developed by Google – part of the inspiration for Hadoop;
- Hadoop - an open source software framework for processing huge datasets on certain kinds of problems on a distributed system;
- Hbase - an open source distributed, non-relational database modeled on Google's Big Table;
- MapReduce - a software framework introduced by Google for processing huge datasets on certain kinds of problems on a distributed system;
- Mashup - an application that uses and combines data presentation or functionality from two or more sources to create new services.

CONCLUSION

This research was motivated by two considerations: (1) Big Data is a component that may offer a new value for organizations, (2) too little research has been conducted worldwide to focus on BD applications. This paper discussed the issue of BD and possibilities that it offers for organizations. It has described and distinguished different BD application areas as well as the various techniques for analyzing BD. They included e.g.: predictive modeling and advanced data mining, web mining, agent-based models, genetic algorithms, neural networks, network analysis, spatial analysis, visualization tools. These techniques help organizations to overcome the bounder of actual knowledge. They are focused on searching of new knowledge sources, enriching of existing resources, adoption of new behavioral orientations and acquisition of new competencies. They are predisposed to support the long term decisions, prediction and optimization as well as to introduce new business models. However, it should be noted that the usage of such BD techniques requires employees of organizations to possess an "analytical erudition", soft skills, e.g., culture based on facts and knowledge, trust, human resources management or managing analytical/creativity people. Therefore our future research will be focused on capabilities and skills that are required to adopt BD in organizations.

The study makes theoretical contribution to the relevant literature. BD issue is generally an unexplored field of research. Therefore, the current study contributes to the existing knowledge on BD by investigating the various BD application areas as well as different techniques for BD analyzing.

REFERENCES

Baaras, H., & Kemper, H.G. (2008). Management support with structured and unstructured data – an integrated Business Intelligence framework. *Information Systems Management*, 25(2), 132-148.

Borgatti, S.P., Everet, M.G., & Freeman, L.C. (2002). *UCInet for Widows: Software for Social Network Analysis.* Harvard. MA: Analytic Technologies.

Capgemini (2013). Consulting. Technology. Outsourcing. Retrieved September 2013 from http://www.capgemini.com/sites/default/files/resource/pdf/Search-Based_BI.pdf.

Chen, H., Chiang, R.H.L, & Storey, V.C. (2012). Business Intelligence and analytics: from Big data to big impact. *MIS Quarterly*, 36(4), 1-24.

Chung, W. Chen, H., & Nunamaker, J.F. (2005). A visual framework for knowledge discovery on the web: An empirical study of business intelligence exploration. *Journal of Management Information Systems*, 21(4), 57-84.

Clavier, P. R., Lotriet, H., & Loggerenberger, J. (2012*). Business Intelligence Challenges in the Context of Goods-and Service-Domain Logic. Proceedings of 45th Hawaii International Conference on System Science*, IEEE Computer Society, 4138-4147.

Davenport ,T.H., & Harris, J.G. (2007). *Competing on Analytics. The New Science on Winning*. Boston Masssachusetts: Harvard Business School Press.

Ferguson M. (2012). *Architecting a big data platform for analysis. Intelligent Business Strategies*, IBM, white paper.

Han, J., Kamber, M., & Pei, J. (2011). *Data Mining. Concept and Techniques*. New York: Morgan Kaufmann.

Howson, C. (2008). *Successful Business Intelligence: Secrets to Making BI a Killer Application*. New York: McGraw-Hill.

Ishikawa, A., Nakagawa, J. (2013). *An Introduction to Knowledge Information Strategy. From Business Intelligence to Knowledge Sciences*. World Scientific, London.

Karim, A.J. (2011). The value of Competitive Business Intelligence System (CBIS) to Stimulate Competitiveness in Global Market. *International Journal of Business and Social Science, Special Issue,* 2(19), 196-203.

Larose D.T. (2005). *Discovering Knowledge in Data. An Introduction to data Mining*. New York: John Wiley & Sons, Inc.

Liautaud, B., & Hammond, M. (2002). *E-Business Intelligence. Turning Information into Knowledge into Profit*. New York: McGraw-Hill.

Manyika, J. Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., Byers, A. H. (2011). *Big data: The next frontier for innovation, competition, and productivity***.** MCKINSEY GLOBAL INSTITUTE.

Maybury, M.T. (2004). *New Direction in Question Answering*. Cambridge: The MIT Press.

Moss, L., & Atre, S. (2003). *Business Intelligence Roadmap: The Complete Lifecycle for Decision-Support Applications.* Boston: Addison-Wesley.

Negash, S., & Gray, P. (2008). *Business Intelligence*. In F. Burstein, C.W. Holsapple (eds.), *Decision Support Systems*. Berlin: Springer, 175-193.

Nelson G.S. (2010). *Business Intelligence 2.0: Are we there yet?*. SAS Global Forum. Business intelligence/Analytics. North Carolina, http://support.sas.com/resources/papers/proceedings10/040-2010.pdf.

Newman, M.E.J. (2010). *Networks. An Introduction*. Oxford: Oxford University Press.

Olszak, C.M. (2013). The Business intelligence-based Organization - new chances and Possibilities. *International Conference on Management, Leadership and Governance*. Bangkok University, Thailand, 241-249.

Olszak, C.M. (2014). An Overview of Information Tools and Technologies for Competitive Intelligence Building: Theoretical Approach. *Issues in Informing Science and Information Technology*, 11, 139-153.

Olszak, C.M., & Bartuś, T. (2013). Multi-Agent Framework for Social Customer Relationship Management Systems. *Issues in Informing Science and Information Technology*, 10, 368-387.

Poole, D., & Mackworth, A. (2010). *Artificial intelligence: Foundations of computational agent*. Cambridge: University Press.

Poul, S., Gautman, N., & Balint, R. (2003). *Preparing and data mining with Microsoft SQL Server 2000 and Analysis Services.* New York: Addison-Wesley.

Robins, G., Pattison, P., Kalish, Y., & Lusher, D. (2007). An Introduction to Exponential Random Graph (p*) Models for Social Networks. *Social Networks,* 29(2), 173-191.

Sauter, V.L. (2010). *Decision Support Systems for Business Intelligence*. New Jersey: Wiley.

Steyl, J. (2012). *Knowledge Management - BI vs. CI.* Retrieved August 2012 from http://it.toolbox.com/blogs/bi-ci/business-intelligence-vs-competitive-intelligence-32441.

Tan, P.N., Steinbach, M., & Kumar, V. (2005). *Introduction do Data Mining.* New York: Addison-Wesley.

Thomsen, E. (2002). Agents uncovered. *Intelligent Enterprise*, 5(15), 45.

Veber, J. (2012). Operational-economic aspects of cloud computing. In M. Tvrdikova, J. Minster, P. Rozenhal (eds.), *IT for Practice*. Ostrava: VSB-TU Economicka Faculta, 178-186.

Vercellis, C. (2009). *Business Intelligence*. Chichester: Wiley.

Weiss, A. (2002). A brief guide to competitive intelligence. *Business Information Review*, 19(2).

Williams, S., & Williams, N. (2007). *The Profit Impact of Business Intelligence*. San Francisco: Morgan Kaufmann.

Witten, I.H., Frank, E., & Hall, M. (2011*). Data Mining: Practical Machine Learning Tools and Techniques.* San Francisco: Morgan Kauflmann.

Wixom, B.H., & Watson, H.J. (2010). The BI-based organization. *International Journal of Business Intelligence Research,* 1, 13-28.