PERSPECTIVES ON DATA ENGINEERING: EMERGING ROLE IN ORGANISATION DEVELOPMENT

Sławomir LUŚCIŃSKI

Summary: The purpose of this article is to present data engineering as an emerging and important element in organisation development. The Big data phenomenon is characterized in relation to new trends in production engineering and data processing for analytical purposes in extended-enterprise environment. An impact of big data opportunities on organisational development as well as challenges, limitations and possible scenario of data engineering role in achieving strategic alignment between business domain and IT domain of organisation are presented.

Keywords: big data, data engineering, management information systems, organisation development

1. Introduction

It is estimated [1] that total amount of digital data, created and replicated both enterprises and consumers all over the world, in the year 2013 reached the level of 4.4 zettabyte (10^{21} bytes). It is expected to grow exponentially to 44 zettabytes in 2020 . Enterprises are responsible for substantial part of this world-wide data ocean – est. 85%. – however they produced only 1.5 zettabyte.

A process of rapid growth of resources of digital data is, among others, driven by:

- development of the microprocessor (1971), which introduce concept of software-driven microelectronics devices that are the core element of digital revolution in consumer and professional electronics, including process of digital convergence of analogy signal processing to digital signal processing devices and development of embedded systems,
- technological revolution that results in development of mass-storage devices, offered both for professional and consumer use, with enormous capacity and decreasing cost-per-byte: from 9,200US$/MB in 1956 to 0.0000350 US$/MB in 2014 [2],
- achievements in mathematics of information and coding,
- achievements in computer science (databases, programming languages etc.),
- development of personal digital devices (personal computers, tablets, smartphones, digital cameras, recorders, GPS, etc.)
- development of global computer network system (internet) that results in arising the Information Society consist of those members of global community, who use growing amount of internet services in everyday life (e-mail, WWW, social services, cloud-computing etc.),
- development of wireless communication solutions (WiFi, RFID, Bluetooth, NFC, GSM etc.) including broadband wireless access to Internet,

– development of the Internet of Things (IoT): intelligent devices with connectivity enablement.

IT revolution brings widespread use of *datafication* processes in every area of human cognition. The *datafication* means more than digitalization; it is the process of putting any phenomenon into a quantified format, enabling analysis for prediction and planning [3]. Modern technologies, like Internet, global positioning systems (GPS, GLONASS), mobile telephony, information systems etc. create enormous and incomparable to nothing before, volumes of data stored in electronic format, which is accessible with use of computers.

2. The big data definition

D. Laney in research note entitled *3-D Data Management: Controlling Data Volume, Velocity and Variety,* published in February 2001, formulated "3V's" framework for dealing with growing wave of data flowing through e-commerce channels [4]. However the term "big data" had emerged as a popular colloquialism [5]. It was finally popularized and defined in McKinsey report [6] as a term which "refers to datasets whose size is beyond the ability of typical database software tools to capture, store, manage, and analyse". The definition presented by Wiki indicates that the term "big data" may be used "for any collection of data sets so large or complex that it becomes difficult to process them using traditional data processing applications" [7]. More descriptive definitions of big data follows the Laney 3-V model (Tab. 1).

Tab. 1. The 3-V dimensions of Big Data

Challenges in data growth	**Range of Values**
Data Volume	Shift from scale of : 1000^2 bytes – Megabyte (KB) in 70', through 1000^3 bytes – Gigabyte (GB) in first decade of XXI century to: 1000^4 bytes – Terabyte (TB); 1000^5 bytes – Petabyte (PB) 1000^6 bytes – Exabyte (EB); 1000^7 bytes – Zettabyte (ZB) 1000^8 bytes – Yottabyte (YB)
Data Velocity	Shift from batch, periodic data processing to: near Real Time or Real Time processing
Data Variety	Shift from numbers structured into tables or databases to: photo images, text (Websites, social media), audio, video, unstructured set of data, logs etc.

Source: own

Data Volume refers to the growing amount of data. It is assumed that typically data volume considered to be "big" starts from terabytes. According to official claims, The Utah Data Center will be the first to reach capacity of 1 yottabytes stored surveillance data [8].

Data Velocity represents speed of new data creation and speed of data access. Variety describes the range of different data sources and types, which can be more or less structured. When we consider big data characteristics in 3-V dimensional approach, we assume that we are dealing with issues which are grounded in big volume of data, enabling high velocity (close to Real Time access) and big variety connected with highly unstructured form.

The fourth V is frequently added do this characteristic of Big Data, which stands for Valuable from the decision-makers point of view. Big data creates value in several ways [6]:

- creating transparency – big data accessible across organisational boundaries in public sector as well as in business enable communication and effectiveness of processes, for instance in relation to reduced time and cost, as well as quality improvement,
- enabling experimentation to discover needs, expose variability and improve performance,
- segmenting populations to customize actions – big data gives the opportunity to better understanding of customer specific needs and more effective tailoring of consumer goods and services,
- replacing or supporting human decision making with automated algorithms,
- innovating new business models, products, and services.

According to the research results, companies in the top third of their industry in the use of data-driven decision making were, on average, 5% more productive and 6% more profitable than their competitors.

3. Managerial decision making challenges

Management of the organization is a continuous decision making process at different levels of the hierarchy and responsibility, which may be grouped into three stages: (1) *Strategic planning* – Senior Managers; (2) *Managerial control* – Middle Managers; (3) *Operational control* - Operational Managers. Decisions are made in relation to various functional areas: Sales and Marketing, Manufacturing and Production, Finance and Accounting, Human Resources [9].

The decision makers act under the pressure of limited resources and time: in today's competitive business environment, managers are facing challenges in dealing with big data issues of rapid decision making. Increase of the amount of available information is associated with an additional consumption of resources and time, but does not lead to a situation in which the importance and meaning of information would be rewarding.

Managers deal with growing uncertainties, called *generalized uncertainty*, which comes with the pace, scope, depth range and speed of propagation of changes in the global economy. The uncertainty in decision making process falls into three categories [10]:

1. Uncertainty of state - unpredictability in relation to an incident in the environment.
2. Uncertainty of effect - unpredictability of results of the event and the associated cause-and-effect relations.
3. Uncertainty of reaction - lack of information on possible options for action in response to the incident and the possible consequences.

Following R.B. Duncan, we may say that *perceived environmental uncertainty* (PEU) can be interpreted as the result of the impact of environment complexity and dynamism [11]. The other explanation of managerial uncertainty is that managers perceive the environment to be unpredictable, and this occurs when the lack information that they feel are needed to undertake decisions.

Managerial decision making challenges can be described as follows [12]:

- managers need to analyse large amounts of information while making a decision, solving a problem, or appraising an opportunity,
- managers must make decisions quickly, with a time constraints, usually based on incomplete information,

– managers must apply sophisticated analysis techniques to make strategic decisions.

Organisation in order to deal with uncertainty grounded in lack of information, develops the Information System (IS). IS is a formal, sociotechnical, organizational system, designed to collect, process, store and distribute information**.** The Information System is formed along with the evolution of the organisation, and in view of the increase in the demand for information arising from the development of the organisation, tends to sprawl.

Management information system (MIS) is the "extracted part of the information system, which from the point of view of the objectives, is computerized" [13]. E. Turban defines MIS as "the formal computer system making the choice, sharing and integration of data from different sources in order to provide in time the necessary information for decision making".

Following J. Senna, information useful in decision-making in the organization falls into two categories: 1) accounting information 2) managing information. Due to this classification, MIS may be divided into two types [14]: 1) evidencing and operative systems, 2) information and analytical systems. The evidencing and operative systems support daily activities of an organization, include a record of events and automate ongoing operations, provide accounting information useful mainly in case of operational and partly on managerial level of responsibility. The Information and analytical systems are focused on analysing and providing information useful in decision-making mainly over the operational management level.

4. The evidencing and operative systems evolution

Generally, and as identified in Fig. 1, there are three types of real-world processes that generate data streams, which may be both considered as Big Data sources separately, as well as in total.

Fig. 1 The nature of the Big data sources. Own, based on [15].

Data observations are generated, among others, by wide range of measuring instruments, GPS systems, surveillance systems etc. Transactions are generated internally in enterprises, as well as in the extended enterprise, that consist of enterprises involved in the Collaborative commerce (C-commerce). A term Collaborative commerce (C-commerce) describes electronically enabled business interactions among an enterprises' internal personnel, business partners and customers throughout a trading community. The trading community could be the whole industry, industry segment, supply chain or supply chain segment [16]. To meet challenges of C-commerce, next-generation ERP systems were developed. Following Gartner (2000) that coined term ERP II, it is "a business strategy and a set of industry-domain-specific applications that build customer and shareholder value by enabling and optimizing enterprise and inter-enterprise collaborative operational and financial processes" [17].

The conceptual framework of ERP II may be seen as four layers structure consist of four layers: 1) Foundation, 2) Process, 3) Analytical and 4) Portal (Tab. 2).

Tab. 2 Layers of ERP II

Layer	Components	
Foundation	Core	Integrated Database (DB) Application Framework (AF)
Process	Central	Enterprise Resource Planning (ERP) Manufacturing Execution Systems (MES) Business Process Management (BPM)
Analytical	Corporate	Supply Chain Management (SCM) Customer Relationship Management (CRM) Supplier Relationship Management (SRM) Product Lifecycle Management (PLM) Employee Lifecycle Management (ELM), Corporate Performance Management (CPM).
Portal	Collaborative	Business-To-Consumer (B2C) Business-To-Employee (B2E) Enterprise Application Integration (EAI)

Source: [18]

Interactions data is mainly fuelled by Internet of Things (IoT) and is the biggest growth part of Big data. These data are mainly supplied by embedded systems. It is estimated that it will growth from 8% in 2014 to 21% in 2020 [1]. Term IoT describes network of physical objects with embedded systems, which are able to interconnected through Internet Protocol (IP), both through wireless and wired way. The IoT produces huge volumes of data from sensors as well as demand for control data (for instance for actuators). More generally, so called Cyber-Physical Systems (CPS) consist of embedded systems (part of devices, buildings, transport routes, vehicles, production facilities etc.) that use Internet services. Networked CPS create distributed, intelligent systems, with connectivity enablement feature, which can be assembled into smart grid for manufacturing purposes. Thanks to networking through the Internet, CPS can use and produce worldwide available data and services. In the near future, it will be possible to employ so called *Function production engineering* into process of development of production environment (production lines); engineers will be able to interact with functionalities of automation system only, broadcasted as a services in networked complex production systems, instead of interacting with abstract signals and complex interfaces.

Fig. 2 presents authors' proposition of modified stage model of Manufacturing planning and control (MPC) systems evolution, originally proposed by Rondeau & Litteral (2001). Modification is grounded in the idea of the 4th industrial revolution (Industry 4.0), driven by advanced information and communication technologies (ICT), which are becoming increasingly prevalent in industrial automation [19, 20].

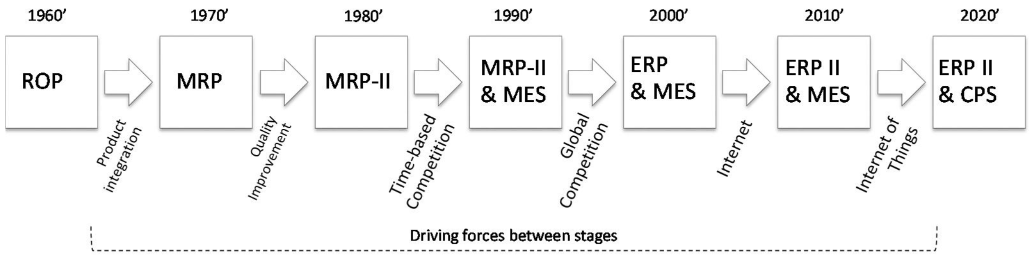

Fig. 2 Manufacturing Planning and Control stage model - modified.
Adapted from: [21]

Manufacturing (or production) planning and control have evolved through five major stages due to forces driving the evolutionary process between stages [21]:

1) reorder point (ROP) systems,
2) materials requirement planning (MRP) systems; due to production integration,
3) manufacturing resource planning (MRP-II) systems; due to quality improvement,
4) MRP-II with manufacturing execution systems (MES) ; due to time-based competition,
5) enterprise resource planning systems (ERP) with MES; due to global competition.

Two additional stages may be added:

6) extended ERP (ERP II) with MES – transition driving force: internet,
7) extended ERP (ERP II) with autonomy and self-control through CPS intelligence (elimination of MES) – transition driving force: internet of things and services.

5. The Information and analytical systems evolution

Simon (1960) explains decision making as sequential process consisting of four stages: (1) intelligence, (2) design, (3) choice, and (4) implementation [9]. Design stage involves identifying and exploring various solutions of the problem. Choice stage is for choosing among solutions alternatives. Implementation stage includes making the chosen alternative work and monitoring achieved results. Intelligence stage consists of discovering, identifying, and understanding the problems occurring in the organization.

Term Business Intelligence (BI) is used in relation to the infrastructure for warehousing, integrating, reporting and analyzing data that comes from business environment. There are six analytic functionalities that BI systems deliver to decision makers [9]: 1) pre-defined production reports, 2) parameterized reports, 3) dashboards and scoreboards, 4) ad-hoc queries and searches, 5) the ability to drill down to detailed views of data, and 6) the ability to model scenarios and create forecasts.

The main components of BI are enterprise data warehouse and data marts. Data warehouse solution architecture (DW) is a structured data repository focused on data analyses needs, which collects and consolidates data from multiple source systems. Data mart (DM) is a scaled-down data warehouse, which meets the needs of specific audience. Data stored in a DW/DM can be seen as facts and as dimensions (reference information associated to the transaction or the event). Creating and maintaining of data warehouse is proceeded with data manipulation tools called ETL (Extraction, Transformation, and Load) - whose functions are as follows [9]:

- extraction of data from outside sources, consolidation and conversion to adequate format,
- data transformation with use of set of rules and functions to derive (from previously extracted data) the data to be loaded: data cleansing, deduplication, encoding, merging, junction, grouping, summarization,
- load of transformed data into a data warehouse.

For the analytic purposes, BI systems employs a software tools that include:

- relational reports for visualising data stored in a DW (tables, graphs),
- On-Line Analytical Processing (OLAP) software to aggregate data from DW into OLAP cubes – database structure optimised for querying and reporting e.g. quick response to ad-hoc queries,
- data mining tools to automatically discovering nonobvious relationships in large databases through identifying significant data patterns.

It is widely recognized, that Big data is becoming a next key basis of competition and growth for firms. Big data analysis seem to be the driving force of new product and service development, as well as creativity and innovativeness. Big data, as a tool, requires new approaches and techniques, which are able to deal with issues that arise from inherent properties of big data resources: big volume of data, as well high velocity and big variety connected with highly unstructured form.

New approach to data analysis is characterized by three rules, which emerged as a new paradigm, which are used in development of new data exploration techniques [3]:

1) collect and use a lot of data rather than restrict inference to a predetermined sample of data (rule n=all),
2) accept imperfect data; "benefits of using vastly more data of variable quality outweigh the costs of using smaller amounts of very exact data", as well as allow to timing decision window of opportunity; there is a trade-off between uncertainty and missed opportunity.
3) using big data frequently means "forgoing the quest for why in return for knowing what" (to do things done); accepting discovered correlations as a sufficient result of proceeded analysis for e.g. prediction purposes.

New techniques employed to big data analysis, that support and fulfil rules presented above, draw from several disciplines, including statistics, applied mathematics, and economics, computer science. They include (among others): cluster analysis, data fusion and integration, data mining, genetic algorithms, machine learning, natural language processing, neural networks, network analysis, signal processing, spatial analysis, simulation, time series analysis, and visualization [6] .

An integrated databases solutions should be empowered through employing such approaches to data management like [4]:

- profiling data sources to identify and eliminate redundancies, monitoring data usage to move out unused data (to deal with data volume),
- designing architectures that balance data latency with decision cycles to meet near real time response demand (to deal with data velocity),
- new integration approaches to relating data of various incompatible types and use of metadata management solutions (to deal with data variety).

The Apache Hadoop open-source software platform (developed by The Apache Software Foundation) for reliable, scalable and distributed computing, became a standard in data processing [22]. There is a variety of market players, which are involved in supplying value added Hadoop distributions or implement Hadoop's core modules in their products. First group includes e.g. Cloudera, and MapR. The other group is consist of companies like: Intel, Microsoft, Oracle, SAP. The Hadoop software enables the distributed processing of large data sets across clusters of computers, with software-supported reliability. The project includes the core modules: Hadoop Common, Hadoop Distributed File System (HDFS), Hadoop YARN, and Hadoop MapReduce. Hadoop YARN is a framework for job scheduling and cluster resource management. Hadoop MapReduce is a software framework for writing applications in Java (supports also Hive or Pig query specification layers) which process vast amounts of data in-parallel on large clusters.

6. The big data impact on organizational development

Henderson and Venkatraman argue [23], that the inability to realize value from IT investments is due to lack of alignment between the business and IT strategies of organisations. They formulated and developed the model termed Strategic Alignment Model (Fig 3). Strategic alignment occurs as a dynamic process between four domains: business strategy, IT strategy, organisational infrastructure and its processes, IS infrastructure and its processes.

Four alignment perspectives are identified as sequences of paired integrations (two-stages process):

1) Strategy Execution Alignment Perspective – business strategy drives organisational infrastructure which subsequently drives IS Infrastructure (functional integration follows strategic integration).
2) Technology Transformation Perspective – Business Strategy drives IT strategy, which in turn drive IS Infrastructure (strategic integration follows functional integration).
3) Competitive Potential Perspective – IT Strategy drives Business Strategy, which in turn drives organisational Infrastructure (strategic integration follows functional integration).
4) Service Level Alignment Perspective – IT Strategy drives IS Infrastructure, which in turn drives Organisational infrastructure (functional integration follows strategic integration.

The big data revolution induces process of changes in organisations, which can be characterised from the Service Level Alignment Perspective. New capabilities of information technology, rooted in accessibility to big data resources as well as to big data techniques and methods (existing in the external domain of organisation) should reflect in tailoring IT Strategy. Implementation of new strategy requires introducing new technological achievements in data processing in relation to evidencing and operative systems as well as to information and analytical systems, as it concerns to stage of strategic integration. Those changes should be followed in stage of functional integration through internalization of new services supporting decision makers in decreasing the level of uncertainty.

The Map of Organization Development describes in more detailed way the dynamic nature of an impact of information systems on organization development in terms of their effective use in the management [24 – 26]. Model enables both diagnose of development state and prognosis of changes in considered domains, with employing defined quantitative indices in order to parameterize those domains and information feedbacks between them. Three *development gaps* of information systems are defined in model to measure discrepancy between the levels of application portfolio abilities, utilization and managerial expectations,

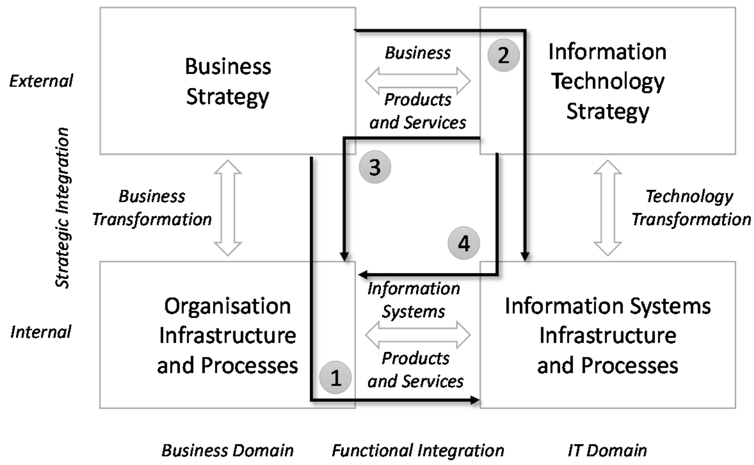

Fig. 3 The Strategic Alignment Model. Adapted from: [23]

as follows:

- *utilization gap* – between the current and a potential way of using held IS;
- *expectations gap* – between the desired and a potential way of using held IS;
- *technology gap* – between the potential and a desired way of using held IS.

The Map of Organization Development model evolves to more detailed description of dynamic in organisation development, through identifying three development gaps in managerial practice that impact organisation readiness and effeteness in process of information systems driven organizational change [27] :

- *coherence gap* - a measure of the lack of coherence (gap) between the formal organizational structure (as declared) and management practices (as is),
- *anticipation gap* – a measure of the discrepancy between the expected (as should be) and the current (as is) level of management practices,
- *implementation gap* – a measure of the discrepancy between the actual formal organizational structure (as declared) and the expected (as should be).

A changes in management of organisation may be also considered in relation to adequate level of internal innovativeness' environment maturity [28].

7. The emerging role of data engineering

Creating a business value from big data opportunities is the multi-stage value chain [29]. There are at least four primary stages in every data value chain (Fig. 4).

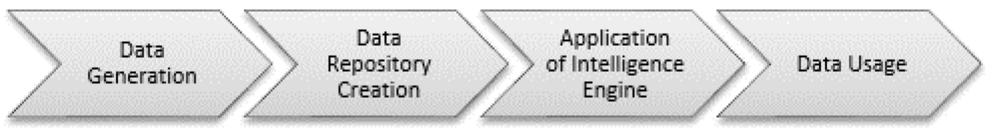

Fig. 4 The big data value chain. Adapted from: [29]

The Data Generation stage includes collecting data where it originates, and making them accessible as row data stream. At the Data Repository Creation, raw data from selected sources are gathered and stored in format that is convenient for analytics purposes. The Application of Intelligence Engine stage includes applying analytics and algorithms to transform raw data to suitable aggregated data. At the final stage, the outputs of the intelligence engine are converted to tangible values. A stages are interdependent and tend to belong to separate entities. Hence there is a continuum of change a proportion of integrity and modularity of data value chain; the level of modularity depends on number of entities involved in data value chain and the number of data transition points between cooperating entities (30).

Assuming that organisation, searching value from big data, is involved in at least one of stages and in at least one of data value chains, we should consider a possible variety of data flowing through, followed by variety of information systems' architecture. It means, that process of managing information as a strategic resource for improving organizational performance should include information systems agility as well as information quality. Growing interest in big data is not limited to information technology itself, but is visible in emerging, relatively new role in IS development that is termed as data engineering.

The goal of the data engineering in big data valuation processes, is to build up a data foundation for analysis purposes (Fig. 5).

The key element of data foundation is a data engineering platform [31], that serves as an unifying platform for collecting, cleaning, integration and storage data. In relation to big data, such data platform should have ability to collect unstructured data, which is supplied from various sources. A data processing on data engineering platform is designed as data pipeline to enable data buffering and enhance parallelized vast amount of data execution (i.e. the Apache Hadoop platform mentioned above).

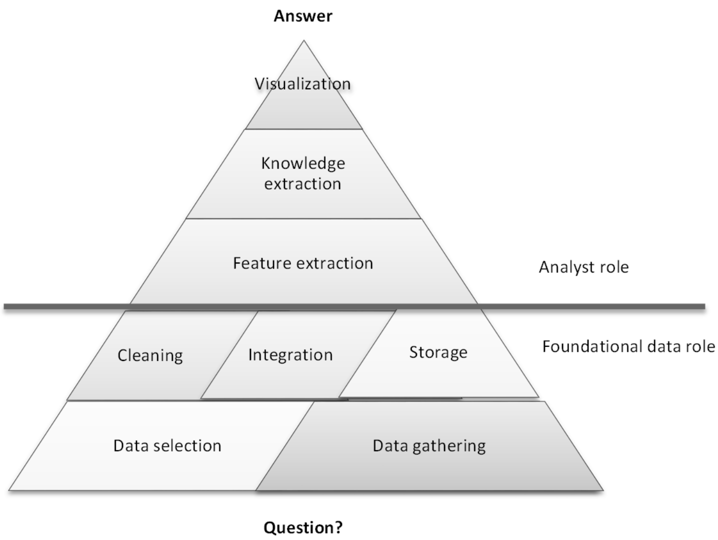

Fig. 5 The pyramid of data analysis. Adapted from: [30]

The distinctive feature of a data engineer in organisation is that it requires performing mainly a foundation data role as well as active supporting for analyst role. From the perspective of The Map of Organization Development [32], data engineering may be crucial in closing the development gaps of both information systems (foundational data role) and management (analyst role).

Following the Technical Committee on Data Engineering (TCDE) of the IEEE Computer Society, data engineering covers [33]:

- data management systems and modern hardware/software platforms,
- data models, data integration, semantics and data quality,
- spatial, temporal, graph, scientific, statistical and multimedia databases,
- data mining, data warehousing, and OLAP,
- big data, streams and clouds,
- information management, distribution, mobility, and the WWW,
- data security, privacy and trust,
- performance, experiments, and analysis of data systems.

8. Conclusions

Technological achievements drives development of Information Society. The inherent feature of development of the Information and Communication Technologies is exponentially growth of gathered data. The big data create a new possibilities in decreasing *generalized uncertainty* through the use of next-generation Business Intelligence systems. However, development processes are not limited to IT itself. An ability to realize value from IT investments in big data is due to achieve (or maintain) an alignment between the business domain and IT domain of organisation. Creating a business value from big data opportunities is linked with challenges grounded in the big data value chain, that brings agility demand to business information systems. A lack of agility in terms of systems architecture may be followed by lack of functional integration between the business domain and IT domain on the internal level of organisation. A possible causes can be identified as development gaps both in management and in information systems.

To avoid obstacles in organisation development and gain the business value from big data opportunities, organisations should enrich their core competences in supplying business information for decision makers through applying data engineering.

Emerging role of data engineering in organisation development creates market demand for new kind of specialist. Education and professional training of data engineers appears to be most actual challenge for educational institutions. There is also an interesting perspectives for future researches in business information management as well as in computer science.

Bibliography

1. EMC Corporation, International Data Corporation (IDC). The EMC Digital Universe study. The Digital Universe and Big Data - EMC . [Online] EMC Corporation, 2014. [Cited: 6 Grudzień 2014.] http://www.emc.com/leadership/digital-universe/index.htm.
2. McCallum, J.C. Disk Drive Prices (1955-2014). Personal Website of John C. McCallum. [Online]14 Sep. 2014 . [Cited: 21 Dec. 2014.] http://www.jcmit.com/diskprice.htm.
3. Mayer-Schönberger V. i Cukier K. BIG DATA. Rewolucja, która zmieni nasze myślenie, pracę i życie. Warszawa : Wydawnictwo MT Biznes, 2014. ISBN: 978-83-7746-515-8.
4. Laney, D. 3-D Data Management: Controlling Data Volume, Velocity and Variety. Doug Laney. A Member of The Gartner Blog Network. [Online] 6 Feb 2001. [Cited: 30 Dec 2014.] http://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf.
5. Holton, L. Britannica - The Online Encyclopedia. Big Data: Mining for Nuggets of Information: Year In Review 2013. [Online] Encyclopædia Britannica, Inc., January

29, 2014. [Cited: Dec. 20, 2014.] http://www.britannica.com/EBchecked/topic/1957644/Big-Data-Mining-for-Nuggets-of-Information-Year-In-Review-2013#toc314304.
6. Manyika, J., et al. Big data: The next frontier for innovation, competition, and productivity. McKinsey&Company Home Page. [Online] June 2011. [Cited: 1 Dec 2014.] http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation.
7. Wikipedia. "Big data". Wikipedia, the free encyclopedia. [Online] Wikimedia Foundation, Inc., 13 January 2015. [Cited: 14 January 2015.] http://en.wikipedia.org/wiki/Big_data.
8. Herbert, G.R. 2012 Energy Summit - Official Blog for Utah Governor Gary R. Herbert. Official blog of the Governor of Utah. [Online] State of Utah, Feb 2012. [Cited: 10 Dec 2014.] http://blog.governor.utah.gov/2012/02/2012-energy-summit/.
9. Laudon, K. C. and Laudon, J. P. Management information systems: managing the digital firm. Twelth edition. Essex : Pearson Education Limited, 2012.
10. Grote, G. Management of Uncertainty. Theory and Application in the Design of Systems and Organizations. London : Springer-Verlag London Ltd., 2009.
11. Hatch, M. J. Teoria organizacji. [trans.] Paweł Łuków. Warszawa : Wydawnictwo Naukowe PWN SA, 2002.
12. Baltzan, P., et al. Business Driven Information Systems. 2e. s.l. : McGraw-Hill Education - Europe, 2009. ISBN: 9780070277274.
13. Kisielnicki, J. and Sroka, H. Systemy informatyczne biznesu. III uzupełnione i zmienione. Warszawa : Wydawnictwo PLACET, 2005.
14. Januszewski, A. Funkcjonalność Informatycznych systemów zarządzania. Warszawa : Wydawnictwo Naukowe PWN, 2008. Vol. 1.
15. Connolly, S. 7 Key Drivers for the Big Data Market. The Hortonworks Blog. [Online] Horton Works, May 14th, 2012. [Cited: September 25th, 2014.] http://hortonworks.com/blog/7-key-drivers-for-the-big-data-market/.
16. Gartner. Term: Collaborative Commerce (C-commerce). Gartner IT-Glossary. [Online] Gartner, Inc., 2013. [Cited: 20 Feb 2013.] http://www.gartner.com/it-glossary/c-commerce-collaborative-commerce.
17. Bond, B., et al. ERP Is Dead — Long Live ERP II. GartnerGroup RAS Services / SPA-12-0420. http://www.pentaprise.de/infocenter.php. [Online] 4 October 2000. [Cited: 19 Czerwiec 2009.] http://www.pentaprise.de/cms_showpdf.php?pdfname=infoc_report.
18. Møller, C. ERP II: a conceptual framework for next-generation enterprise systems? Journal of Enterprise Information Management. 2005, Vol. Vol. 18, No. 4, pp. p. 483-497.
19. Actech Communication Promoters Group of the Industry-Science. Recommendations for implementing the strategic initiative INDUSTRIE 4.0. Final report of the Industrie 4.0 Working Group. Website acatech.de. [Online] acatech – National Academy of Science and Engineering, April 2013. [Cited: 5 Jan. 2015.] http://www.acatech.de/fileadmin/user_upload/Baumstruktur_nach_Website/Acatech/root/de/Material_fuer_Sonderseiten/Industrie_4.0/Final_report__Industrie_4.0_accessible.pdf.
20. ROI Management Consulting. Industry 4.0: Future Trend Scenarios through to 2034. http://www.roi-international.com/management-consulting/dialogue/press/singlenews/article/entwicklungsszenarien-fuer-industrie-40-bis-2034.html. [Online] 6 Feb 2014. [Cited: 3 January 2015.] http://www.roi-international.com/uploads/media/PR_ROI_Industry__4__0_MES_0206_2014.pdf.

21. Rondeau, P.J. and Litteral, L.A. The evolution of manufacturing planning and control systems: From reorder point to enterprise resource planning. Production and Inventory Management Journal. 2001, Vol. Vol. 34, No. 2, pp. pp. 1-7.
22. The Apache Software Foundation. Welcome to Apache™ Hadoop®! The Apache Software Foundation. [Online] 12 Dec 2014. [Cited: 17 Jan 2015.] http://hadoop.apache.org/.
23. Henderson J. i Venkatraman N. Strategic alignment: Leveraging information technology for transforming organizations. IBM Sysytems Journal. 1993, Tom 32, 1, strony 4-16.
24. Luściński, S. and Gierulski,W. Mapa rozwoju organizacji. [ed.] Peszko A. Ekonomia-Informatyka-Zarządzanie. Kraków : Wydział Zarządzania AGH, 2002, Vols. Tom III, Zarządzanie.
25. Luściński, S. and Gierulski,W. Model rozwoju i wykorzystania systemów informatycznych zarządzania. [ed.] Knosala R. Komputerowo Zintegrowane Zarządzanie. Warszawa : Wydawnictwa Naukowo-Techniczne, 2010, Vol. II.
26. Luściński, S. Wpływ systemów informatycznych na rozwój organizacji (eng. The Impact of Information Systems on Organization Development). Unpublished doctoral thesis : The University of Warsaw, 2011.
27. Luściński, S. Koncepcja luk w obszarze zarządzania modelu mapy rozwoju organizacji. [ed.] R. Knosala. Innowacje w zarządzaniu i inżynierii produkcji. Opole : Oficyna Wydawnicza Polskiego Towarzystwa Zarządzania Produkcją, 2013, pp. 330-341.
28. Kaczmarska, B. Modelowanie Innowacyjnego Rozwoju Przedsiębiorstw. Kielce : Wydawnictwo Politechniki Świętokrzyskiej, 2015.
29. Gustafson, T. Fink, D. Winning within the data value chain. Strategy & Innovation newsletter, August 2013, Volume 11, Number 2. [Online] [Cited: 10 January 2015.] http://www.innosight.com/innovation-resources/strategy-innovation/winning-within-the-data-value-chain.cfm.
30. Sink D.S. Finding solutions with imperfect information. Industrial Engineer. July 2014, Tom 46, 7, strony 34-39. MasterFILE Premier, EBSCOhost.
31. Nnebedum, V.I, Kamalu, A.U and Dike, J.N. From Data Management to Data Engineering. International Journal of Computer Applications. November 2013, Vol. 81, 11, pp. 36-39.
32. Luściński, S. and Gierulski, W. Diagnose of the organization development - research results. [ed.] R. Knosala. Innovations in Management and Production Engineering. Opole : Oficyna Wydawnicza Polskiego Towarzystwa Zarządzania Produkcją, 2012, pp. 160-174.
33. IEEE TCDE. IEEE Technical Committee on Data Engineering. IEEE Computer Society. [Online] [Cited: 12 January 2015.] http://tab.computer.org/tcde/.

Dr Sławomir LUŚCIŃSKI, MSE
Department of Production Engineering
Faculty of Management and Computer Modelling
Kielce University of Technology
25-314 Kielce, Al. Tysiąclecia Państwa Polskiego 7
Tel. (41) 34-24-378
e-mail: slawomir.luscinski@tu.kielce.pl
Web of Science ResearcherID: C-9015-2012