BIG DATA BASED SYSTEM MODEL OF ELECTRONIC COMMERCE

G. Ilieva*, T. Yankova, S. Klisarova

Faculty of Economic and Social Sciences, Plovdiv University "Paisii Hilendarski", Plovdiv, Bulgaria

ABSTRACT

The purpose of the paper is to systematically analyze the capabilities of big data technology to solve important problems in electronic commerce. The role of big data as one of the central priorities in modern information technologies for online sellers and buyers is underlined. Big data methods to improve the basic functional areas of e-commerce: marketing, payments, supply chain, and management are described. A new model for automation of key business processes in e-commerce through new technology is presented. The model is verified using practical examples of e-commerce activities. The process of e-commerce modernization requires additional infrastructure, new methods and modern software for better data organizing, customer personalization and improved decision making.

Key words: electronic markets, electronic sales, online customer experience.

INTRODUCTION

There is an increasing interest in big data technology, driven by the constant growth of information in electronic commerce companies. Gathered customer data (names, addresses, preferences, etc.) are important firm's assets, but their utilization becomes a more complex and expensive task by the day. On one hand, data volume is so large that its storage and processing using standard programs and devices are very difficult. On the other hand, a significant part of these data is unstructured and their collecting is not easy, but they are exceedingly valuable (for example, likes, tweets, clicks).

Nowadays e-commerce companies need to analyze a large volume and a great variety of data on products, customers, transactions, and deliveries to increase conversion rate. Companies are aware that scale matters and take into consideration many other implications related to successful implementation of big data technology. Big data applications in e-commerce create new opportunities for quick personalization of products and services according to customers' needs, improve the process of integration between partners' value chains and finally, dramatically reduce costs.

Correspondence to: Galina Ilieva, Faculty of Economic and Social Sciences, Plovdiv University "Paisii Hilendarski", Plovdiv, Bulgaria, e-mail: galili@uni-plovdiv.bg,

The paper is structured as follows:

Section 2 introduces in brief the importance of big data technology in e-commerce. In this section, the terms *big data* and big data *business analyses* are defined. The goal and purpose of this paper – to create a new model of big data application in the main functional areas of e-commerce is presented in details. **Section 3** provides an in-depth review of existing methods for solving main e-commerce tasks. Various big data applications in e-commerce were analyzed and they motivated the creation of the functional model from the next section. **Section 4** contains the sample model for big data application in e-commerce, which was developed by the authors and comprises the core business processes in an e-commerce company. The model is described in details, and specific success factors are derived.

The last section is a summary and it offers future work for model validation.

BIG DATA - BASIC CONCEPTS

In the last few years, big data are the subject of a great amount of research. Some research papers delve into the basic concepts, peculiarities, value, development and challenges of big data (1-3), while others analyze the characteristics of big data business intelligence (BI) and business analytics (BA) systems (4-5) by areas of application.

Regardless of the huge interest in the topic, a strict definition of the term *big data* is yet to be

agreed upon. Potential customers try to understand this concept better, in order to be able to reap the benefits of big data application (6). The term *big data*, at first glance, might imply that it relates to management and analysis of large volumes of data. According to one of the most quoted definitions in relevant literature, this term refers to a set of data, whose volume exceeds the capabilities of classical databases to collect, store, manage and analyze information. The term *big data* itself shows that its "definition is intentionally subjective and incorporates a moving definition of how big a dataset needs to be in order to be considered big data." (7). However, big data processing is only a part of the problem. In 2001, Laney described the "3Vs" — volume, variety, and velocity — as the key "data management challenges" (8). In addition to large volume, two more aspects were included: the need to process information with high velocity and a variety of data structures. The last aspect of the term refers to the diversity and lack of structure or weak-structured data. Operating with media content, blog/vlog data, and unstructured documents is becoming increasingly common. Popularizing the term big data, in 2008, Bryant et al. predict that "big data computing" will "transform the activities of companies, scientific researchers, medical practitioners, and our nation's defense and intelligence operations" (9).

Recently, *big data* has started to imply a number of approaches, tools and methods for processing large volumes of structured and unstructured, varied and always ever growing data. Often, behind applications working with big data, there is a distributed computing system, as processing big data requires a set of high productivity machines often combined in a cluster. Big data BI and BA software in e-commerce perform analysis on data generated by transaction systems or other data sources aimed at aiding management decision making. The first challenge here is due to the fact that data are collected from different locations and different systems. Since speed is among the main issues of online analysis, classic data storage is impractical. The second challenge stems from unstructured data which can be stored only on specific platforms. Moreover, unlike classic BI and BA, big data BI and BA are concerned not with descriptive analysis process, but instead with forecasting events before they occur and making recommendations about future actions.

Despite a substantial number of publications devoted to clarifying big data technology, a number of review studies note that, currently, very few empirical studies assessing the true potential of big data in e-commerce have been conducted (6). The major part of papers on e-commerce deals with only a single aspect of e-commerce companies' activities. For example, using data-driven marketing solutions, organizations achieve competitive advantage by successfully overcoming data proliferation and channel fragmentation (10). A rich link between e-commerce and online marketing allows big data analysis to bring additional value to the trading process and its participants (11). The identification of specific patterns of behaviour could also be used to derive precise information about users, such as their current mood and state of mind, as well as their routine behaviour or current affairs (12). Processing large amounts of data provided by internet users can help online operators better understand the preferences and behaviours of their user base (13).

The main purpose of the current paper is researching the role of big data in enhancing the main functional areas of e-commerce: marketing, payments, supply chain and management.

That is the reason the **first task** of this research is making a systematic review and analysis of the many capabilities of big data for automating value chain activities in e-commerce.

Its second task is proposing and verifying a complete functional model for the realization of the main business processes in e-commerce applying big data technology.

BIG DATA TECHNOLOGIES IN E-COMMERCE – STATE OF THE ART

Today, any electronic commerce company, regardless of its size and business processes can take advantage of the potential of large data volumes, accessible from outside sources and stored in its own databases. Companies can now observe its customers, suppliers and competitors in real time.

Due to its large operative and strategic potential, especially in generating business value, big data can transform the whole business model (7). For e-commerce companies this change affects their main functional business model components: marketing, pricing, supply chain and management.

1. Electronic marketing with big data
Electronic marketing is "use of electronic data and applications for planning and execution of the concept, distribution, promotion, and pricing of ideas, products and

services for the creation of exchange that satisfies individual and company goals" (14). With the development of the Internet, there has been a trend of increase in the efforts that enterprises invest in online activities for achieving marketing goals. It is becoming a more frequent occurrence that e-commerce companies use big data for solving e-marketing tasks: they study clients' needs and preferences; advertisements, pricing, etc.

An innovation in online retail is the so-called "smart shops" (shops with digital screens and adaptive window displays), which generate additional internet traffic and sales. In this way, the shopping process transcends the boundary of the physical world and enters the Net. A new retail method arises, what is known as reciprocal selling that combines online and offline elements for stimulating sales (15).
By means of big data, electronic sales, mobile commerce and physical sales are gradually being replaced by offers combining online and offline elements. Retail is turning into total retail.

Customer service is a crucial success factor of e-commerce sites. Recommender systems assist e-commerce providers in aggregating user profiles and making suggestions tailored to user interests derived from collected data. A new large-scale e-commerce recommendation algorithm combines an offline phase with an online one and surpasses traditional similar systems in accuracy and scalability of recommendations (16). Taking into account the differences between e-commerce and traditional retail industry in (17), RFMA clustering model that achieve client's "precision" division was constructed.

For better serving their customers and improving their market presence, companies apply big data text analytics and segment analysis. These are linguistic and statistical techniques (including machine learning), which model and structure the content of text sources for finding valuable information. For example, via text and/or sentiment analysis, unstructured comments by customers on a company's Facebook page or on Twitter can be transformed into valuable information. Socialmention.com and opinioncrawl.com are some examples of frequently used websites for free sentiment analysis.

Personalization of products and services is among the most important applications of big data in e-marketing. Research by Harvard Business Review shows that personalization can increase the ROI on marketing investment between five and eight times and sales by at least 10% (18). Big data are a great opportunity for e-sellers to track user behavior and determine the most effective way of transforming them into loyal buyers (and increasing repeated purchases). Customers shop in a different manner. Data on the time spent at a e-shop by customers are collected in real time, so that a personalized offer (content and/or special sale) can be made. For example, loyal and new customers could be treated in different ways. Personalized offers need to be attractive to new customers and also reward loyal ones.

In addition to crafting special offers for products and services, an important aspect of big data is that they facilitate pricing personalization, also known as **dynamic pricing.** The creation of tailored products and services enables price personalization, too. This means that the determination of a "correct" price of a deal takes into consideration data from multiple sources - competitor pricing, product sales, regional preferences, and customer actions. It has been proven that under some conditions, price discrimination can be mutually advantageous to both firms and consumers (19).

2. Big Data in Electronic Payments

Payments in e-commerce are made via electronic payment systems – the online analogue of traditional payment instruments. The most commonly encountered problems here are related to varios types of fraud: credit card fraud, return fraud and identity fraud.

In the first type, fraudsters use a counterfeit or stolen credit card. This is the most common of the three types. In a return fraud, a product is returned after having been used or it is claimed that the product was never delivered. What is characteristic of this case is that after a consumer's personal information has been stolen, the goods are shipped to a different location.

Using big data to check ordering patterns of each customer, their frequency of ordering, and their order shipping addresses is important to reducing the possibility of fraud. The fraudulent orders characteristic matrix which compares normal and abnormal orders together with a logistic regression model could be successfully implemented to identify frauds (20). Thus, big data and relevant infrastructure successfully enable real time **fraud detection**. Reducing fraud makes a business environment more secure and increases e-commerce profitability.

3. Big Data in Supply Chain

The major supply chain issues in e-commerce pertain to the optimization of non-digital goods delivery and reverse logistics through effective management of flows of material, products and information resources. A good example of supply chain management with big data is predicting the revenue from a certain product in the next quarter. Knowing inventory forecast, a merchant can better manage its inventory costs and avoid key out-of-stock products.

In addition to forecasting inventories, big data enhances **supply chain visibility**. Customers know the exact availability, status, and location of their orders. This can get complicated for retailers if multiple third parties are involved in the supply chain. A customer who has purchased a backordered product would want to know the status. This requires (downstream) communication between warehousing, transportation and with any third-party systems in supply chain.

Using big data, e-commerce companies can build efficient supply chain systems to manage the process of interactions between vendors, logistics, warehousing, delivery, returns etc. Internet of things, RFID, and other new technologies collect and communicate data on a wide range of conditions and redefine supply chain intelligence for better delivery management, improved order picking, facilitated vendor management and automated product sourcing.

4. Management

In global competition, a lot of companies promote and sell their products and services via a web site, online marketing and presence in online social networks. E-commerce management controls firm's e-reputation in accordance with best practices and current technical innovations. The e-commerce company manager, together with the product and marketing managers, are responsible for strategic and tactical decisions for online business development.

The steps of e-commerce decision making process are well known from management theory:
1/ Determining the problem in an unsatisfactory situation;
2/ Examining and defining the problem;
3/ Seeking and analyzing possible alternatives for action;
4/ Choosing among alternatives (decision making);
5/ Implementing the solution;

6/ Assessing and controlling of the result of the solution.

It is evident that big data could improve practically every step of the process, except the realization of the solution. For example, making a decision which products a company should sell online, at what price and aimed at what target market are hard problems for multi-criterion optimization. In order for an acceptable solution to be found, a huge volume of structured and unstructured data needs to be processed. In this case, big data help managers make more informed decisions. Through **data-driven decisions**, a company manages to prevent unnecessary costs, expensive attempts and even wrong decisions.

In marketing, by employing site metrics (total numbers of site visitors, visitors by product or page, etc.), managers work out a marketing plan to increase sales. They analyze transactions data, in order to trace marketing companies' ROI, market channel effectiveness and customer loyalty.

Security is a critical success factor in e-commerce. Together with payment systems vendors, managers develop secure payment mechanisms protecting personal customer data. They are responsible for the reliable storage and protecting customer data from cyber-threats. Big data can be used to ensure a high level of web site security by preventing attacks which would disrupt customers' access to the company's services.

Management in electronic commerce is engaged with predicting future and prospective trends, finding hidden opportunities and planning business development. Applying big data analytics in e-commerce management increases the effectiveness of management activities – planning, organizing, coordinating and controlling due to taking advantage of data-driven decisions.

BIG DATA SYSTEM MODEL OF E-COMMERCE

Many publications describe the business activities in e-commerce conducted with big data, such as advertising, supply chain management and transaction processing (10-11, 13, 16, 21). Few sources, however, fit these activities into a unified whole. The literature provides some guidance for a functional model of e-commerce systems. E-commerce functions that are commonly discussed in the literature include marketing, logistics, electronic payment, management, etc. (22, Ilieva).

E-commerce systems, as information systems, provide information to support all business processes in an organization. Much of this information results from the analysis of data gathered about customers' behaviour. We need to include the necessary functionality in an e-commerce system model to gather and analyse this data to provide information to all functional areas.

Based on our review of the literature and our examination of the questions not addressed in the literature, we propose an e-commerce system model that consists of five interrelated components: electronic marketing, payment, supply chain, infrastructure and management. The central points of the proposed schema are data acquisition and data analysis. How these components interact with each other and with the system users is shown in **Figure 1**.
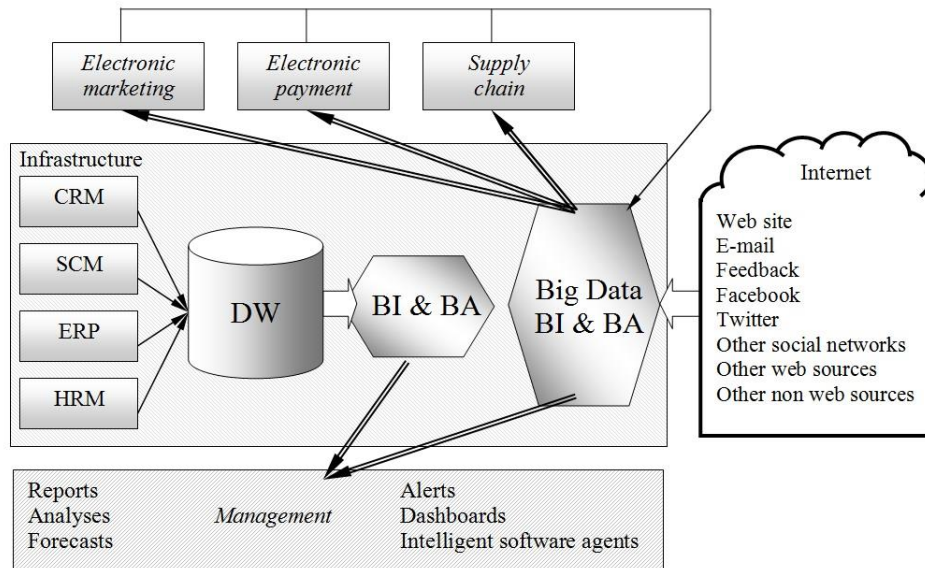

Figure 1. Big data system model of electronic commerce /based on (21)/

We next describe each of the components in **Figure 1.**

The three blocks in the upper part correspond to the three functional areas: e-marketing, e-payments, and supply chain, described in section 3.

Electronic marketing encompasses all marketing activities, performed in e-commerce companies: products and product policy, consumer behavior, advertisement on the Internet, marketing research, etc.

Electronic payments may include credit card, debit card, COD, check (before delivery), and invoice (after delivery). In B2B transactions, electronic funds transfer (EFT) may be a payment option. Usually, this function is linked to an external payment authorization system. It may also be linked to the enterprise's financial accounting or ERP systems.

Supply chain comprises inventory management for timely delivery to customers (product availability) to meet different product and channel demands; minimization of the cost of customer returns and plan for supply-chain needs of the future.

Business intelligence function analyses the data in the data warehouse. Queries, reports and data mining techniques are typically used for this purpose in an effort to identify trends, relationships, and other useful information. The results of such analysis can be used by management for decision making in functional areas, especially marketing.

Big data analysis offline and online.
Big data analytic methods, Real-time analysis and Offline analysis: is usually used for applications without high requirements on response time, e.g., machine learning, statistical analysis, and recommendation algorithms. Big data applications are: text, web, multimedia, network analysis Early Warning, Real-time Monitoring and Real-time Feedback.

Data collection is utilization of special data collection techniques in order to acquire raw data from a specific data generation environment. Two common data collection methods are: log files and network data – company website, online social networks, etc.
Existing storage mechanisms of big data may be classified into three bottom-up levels: (i) file systems, (ii) databases, and (iii) programming models.

New model verification via analysis of existing e-commerce systems

Walmart implemented big data technologies in e-commerce as early as 2011. Its big data department at WalmartLabs developed a goods search engine and during the first year of its use in Walmart.com, its sales grew with 20%. The same technology group developed a software application which had the ability to search and analyze social media applications (like Twitter or Facebook) in real-time in order to provide personalized insights to users.

In order to make the buying process as simple and close to customers as possible, eBay introduced the concept of ZEC (zero effort commerce). Through big data, the system can anticipate what the customer is looking for and save their time and effort in choosing a product. In eBay's "smart shops" in New York and San Francisco choosing clothes is aided by RFID stickers which can also be used for recommendations for accessories that go with the selected product. Anything that has been of interest to a customer is recorded in their profile and later on they could buy the product online. The company has established a Innovation and New Ventures Group which integrates the appropriate e-commerce interfaces in "smart" devices such as smart watches.

Netflix is a classic example of **personalized service**. Using big data, this company registers virtually every customer's interaction with its website. For example, if a customer complains against the online shop's service by filling the proper form or tweeting about it, this information is visible to its call center employees.

For example, Burberry.com, the online store, collects various qualitative data such as browsing patterns, login counts, purchase history, etc. Each client is given a **personalized offer**, which is most relevant to their preferences and behavior. This way, Burberry's saves additional resources by not offering unnecessary discounts to customers that do not need incentives for conversion.

Amazon was among the first companies to implement **dynamic pricing**. E-commerce companies' ability to constantly change pricing on millions of SKUs on a daily basis based on competition, demand for products etc., is a significant source of competitive advantage for them.

PayPal and Amazon have developed **fraud detection** tools that rely on massive datasets containing not only financial details for transactions, but also IP addresses, browser information, and other technical data that will help these companies refine models to predict, identify, and prevent fraud.

For their **sales forecasting**, eBay employ SAP HANA for Early Signal Detection with machine learning algorithms. Wal-Mart analysis show that in poor districts in times of economic crisis, people tend to buy smaller packages. This conclusion was used by the company to plan its inventories. However, after adding more data sources, including those from the company's online store, it was determined that demand for bigger packages increases again during a cold winter, because their customers are stocking up.

In 2014, Amazon got a patent to **ship goods** before customers have even made a decision to buy them, purely based on their predictive big data analytics. Packaged orders can wait in transportation centers, until the relevant order comes. After that the product could be delivered to the customer's address much faster. This system is called Speculative Shipping Process.

CONCLUSION

In recent years, application of big data in managing various aspects of electronic commerce is increasing dramatically. The proposed system e-commerce model based on big data has the potential of enhancing the value chain management of e-stores, optimize marketing activities and increase sales via analysis of consumer behavior, payments, deliveries and competitors. A more accurate assessment of the quantifiable benefits of the proposed model requires processing of empirical data and will be the subject of future research.

REFERENCES
1. Chen, M., Mao, S. and Liu, Y. Big data: a survey. *Mobile Networks and Applications*. 19(2):171-209, 2014.
2. Iafrate, F. A journey from big data to smart data. Digital Enterprise Design & Management. *Advances in Intelligent Systems and Computing*. 261:25-33, 2014.
3. Kaisler, S., Armour, F., Espinosa, A. and Money, W. Big data: issues and challenges moving forward. *Proceedings of 46-th Hawaii International Conference on System*

Sciences (HICSS), pp. 995-1004), 2013. doi:10.1109/HICSS.2013.645.

4. Chen, C. and Zhang, C.-Y. Data-intensive applications, challenges, techniques and technologies: a survey on big data. *Information Sciences*. 275:314-347, 2014.

5. Chen, H., Chiang, R., and Storey, V. Business intelligence and analytics: from big data to big impact, *MIS Quarterly*, 36(4):1165-1188, 2012.

6. Wamba, S., Akter, S., Edwards, A., Chopin, G. and Gnanzou, D. How 'big data' can make big impact: findings from a systematic review and a longitudinal case study. *International Journal of Production Economics*, 165:234-246, 2015. doi:10.1016/j.ijpe.2014.12.031

7. Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., and Byers, A. Big data: the next frontier for innovation, competition, and productivity. *McKinsey Global Institute*, (Online). Available: (2015-07-31). http://www.mckinsey.com/insights/mgi/research/technology_and_innovation/big_data_the_next_frontier_for_innovation

8. Bryant, R., Katz, R. and Lazowska, E. Big-data computing: creating revolutionary breakthroughs in commerce, science, and society. Computing Research Consortium, 2008, Available: http://www.cra.org/ccc/resources/ccc-led-white-papers/ (2015-07-31).

9. Arthur, L. Big data marketing: engage your customers more effectively and drive value, *Wiley*, Hoboken, NJ, USA, 2013.

1. 8. Laney, D. 3D data management: controlling data volume, velocity and variety. Application Delivery Strategies, File 949, *Meta Group*, Stamford, CT, USA, 2001.

10. Pabedinskaite, A., Davidaviciene, V. and Milišauskas, P. Big data driven e-commerce marketing. *Proceedings of 8-th International Scientific Conference Business and Management, Lithuania*, p. 645-654, 2014. doi:10.3846/bm.2014.079

11. Bughin, J., Livingston, J. and Marwaha, S. Seizing the potential of 'big data'. *McKinsey Quarterly*, December 2011, p. 103-109.

12. LaValle, S., Lesser, E., Shockley, R., Hopkins, M. and Kruschwitz, N. Big data, analytics and the path from insights to value. *MIT Sloan Management Review*, 52(2):21-31, 2011.

13. Strauss, J. and Frost, R. E-marketing. *Prentice-Hall*, Englewood Cliffs, NJ, USA, 2001.

14. Apiletti, D. and Forno, F., Desidoo, a big-data application to join the online and real-world marketplaces. New Trends in Databases and Information Systems. *Advances in Intelligent Systems and Computing*. 241:169-175, 2014.

15. Hu, L., Lin, K., Hassan, M., et al. CFSF: On cloud-based recommendation for large-scale e-commerce. *Mobile Networks and Applications*, 20(3):380-390, 2015.

16. Zhang, J. and Zhu, J. Research intelligent precision marketing of e-commerce based on the big data. *Journal of Management and Strategy*. 5(1):33-38, 2014. doi:10.5430/jms.v5n1p33

17. Leibowitz, J., Ungerman, K. and Masri, M. Know your customers wherever they are. *Harvard Business Review*, 26.10.2012. https://hbr.org/2012/10/know-your-customers-wherever-t

18. Rayna, T., Darligton, J. and Striukova, L. Pricing music using personal data: mutually advantageous first-degree price discrimination. *Electronic Markets*. 25(2):139-154, 2015. doi:10.1007/s12525-014-0165-7

19. Yang, Q., Hu, X., Cheng, Z., Miao, K. and Zheng, X. Based big data analysis of fraud detection for online transaction orders. *Cloud Computing Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, 142:98-106, 2015.

20. Nickerson, R. (2002). An e-commerce system model. *Proceedings of the Eight Americas Conference on Information Systems*, p. 310-316, Dallas, TX, USA, 2002.

21. Ilieva, G. Electronic commerce. *Plovdiv University Press*, Plovdiv, Bulgaria, 2010. In Bulgarian