INTEGRATING OPEN DATA INTO COMPANIES' BUSINESS MODELS FOR FOSTERING DIGITAL TRANSFORMATION

Antonova A.[1], Yordanova S.[2],
Sofia University, Bulgaria[1]
University of National and World Economy, Bulgaria[2]

a_antonova@fmi.uni-sofia.bg

syordanova@unwe.bg

Abstract: *Horizontal and vertical integration of companies within the value networks plays a substantial role for faster and wider implementation of Industry 4.0 and adoption of new digital business models. The policy for open data encourages companies to integrate more data flows coming from public sources into their operations. Thus, digital transformation of companies needs to reconsider the available public and open data flows and to improve their business models. The present paper aims to present and analyze how companies can integrate Open data in digital transformation process and new business models` adoption. The main elements of digital business models within Industry 4.0 and smart factories is introduced and discussed. Open data models and standards are analyzed and assessed as a source for value creation within companies. Finally, the paper identifies and discuss the main approaches for companies to implement open data into their business models.*

Keywords: OPEN DATA, BUSINESS MODELS, INDUSTRY 4.0;

1. Introduction

Open Data (OD) and Open Government Data (OGD) became part of large national and international policy initiatives targeting to enforce public authorities to provide open access into public data. Open data became part of government efforts to promote transparency, participation, efficiency and effectiveness in the public sector [Huijboom, Van den Broek, 2011]. Open data policies have been established with the purpose to bring new economic opportunities through encouraging innovations, fostering digitalization, developing advanced products and services and providing further benefits for individuals, organizations and civil society as a whole. As discussed in the report of Carrara et al. (2015), the open data re-use can lead to many political and social benefits as well as to direct and indirect economic gains in the form of new revenues, value-adding activities, costs savings and jobs. Following the general Open data trend, different public, NGOs and private organizations and international bodies, deliver open cross-disciplinary datasets, available for further re-use, naming as for example: www.kaggle.com; google cloud platform (https://cloud.google.com/public-datasets/), open data network (https://www.opendatanetwork.com/) and many others.

However, the public statistics show that the free access to open data is not automatically leading to data use and re-use especially in business context [Welle Donker & van Loenen, 2017]. Even more, unanswered questions remain about the micro-economic and macro-economic effects from Open data re-use and further distribution of the Open Data economic gains [Davis&Perini, 2016]. Moreover, the development of new products and services out of public data and the emergence of new business models prove to be challenging tasks for all stakeholders, including business and public organizations, start-ups and non-for-profits. Even in the most advanced business case scenarios, it becomes clear that further efforts should be made to integrate open data into value-adding business models, business processes and services.

The emerging Industry 4.0 technology paradigm imposes new economy models based on "smart manufacturing" and "smart factories". Organized around big data, connected technologies such as complex cyber-physical systems, Internet of Things, 3D printing, cloud computing, artificial intelligence and robotics have the potential to disrupt value creation along different industries. Big data and advanced data analytics become the key ingredient of the new business models and business processes. However, Sommer (2015) warns, that new Industry 4.0 revolution can change the structure of the business landscape, leaving small companies incapable to survive the digital transformation, unable to adapt to the new-coming business realms. As a result, a question arises: Can thus open data enhance the digital transformation of the SME companies?

The present research aims to outline the opportunities of companies and especially SMEs to integrate open data into their digital transformation process. By exploring the main elements of data-driven value creation process and data-based digital business models, a framework is proposed for value-adding activities of both open data providers and intermediaries within an open data ecosystem.

The paper is structured in four sections. The first section presents the main concepts of open data and the elements of the digital business models transformation within Industry 4.0 and smart factories. In the second section, the data-value driven models for value creation, the data-driven business models and open data value creation is explored. The next section outlines open data models and standards, further making assessment of Open data development in Bulgaria. Finally, the paper identifies and discusses the main difficulties and barriers for companies to implement open data analysis into their business models and there is proposed and discussed a new open data business model, fostering value creation within companies.

2. Open data, Industry 4.0 and company digital transformation

Open data is commonly defined as data, that can be freely used, re-used and redistributed by anyone – subject only at most, to the requirement to attribute and share-alike [Open Data Handbook, Open Knowledge International, 2015]. In order to be reusable, the open data should respond on two main conditions: to be legally open, or released under an open license, legally permitting its re-use and redistribution and to be technically open, or available in an open and machine-readable format, as a complete dataset and preferably for a free download (Open Data EU portal, https://data.europa.eu/).

The open data portals represent a critical data infrastructure as they connect data holders with data users, who in turn can re-use it into products and services that citizens and businesses can benefit from. The main European Open data hub is the European data portal (https://www.europeandataportal.eu/), launched in November 2014 with the aim to facilitate both data publishing and data re-use across different EU countries. The European data portal harvest the metadata of Public Sector Information (PSI) available on public data and geospatial portals across European countries and currently provide access to about 788,671 datasets. Furthermore, the

European data portal provides a wide number of services such as Open data e-learning and training, research publications, public surveys, Open Data barometers, policy recommendations and others. This way, European data portal becomes the main reference center for European open data initiatives. On European scale many other specialized open data portals, such as: INSPIRE (spatial data sets and spatial data services), Copernicus (land, marine, atmosphere, climate change, emergency management and security), GEOSS (Global Earth Observation System of Systems) have been launched. During the next years it is expected the European Open Science Cloud (EOSC initiative) to be launched. All these portals and initiatives come to show that further public datasets will be available for businesses, increasing their potential and capacity to monetize data resources into innovative products and services.

Industry 4.0 and smart factories designate a new trend toward existing industry digital transformation with extended smart and connected technologies. Almada-Lobo (2016) defines Industry 4.0 manufacturing transformation based on Cyber-Physical Systems (CPS) and Cyber Physical Production Systems (CPPS) toward mass customized, decentralized, vertically integrated, connected and mobile, cloud computing and advanced data analysis approach. Saldivar et al., (2015) further summarize that Industry 4.0 leads to a paradigm shift from centralized to decentralized manufacturing based on customer-triggered autonomous processes of cyber-physical systems. Recognizing the economic potential of Industry 4.0, a number of policy strategies have been implemented both on European and on national level [Digitizing European Industry EC portal].

Adoption of Big data or advanced data analysis will play a crucial role in the digital transformation process. As outlined in the whitepaper of Fraunhofer' Institute [Otto et al., 2016], data has changed its economic role in industry from "data as a process outcome" on "data as a process enabler" to become „data as product enabler" and finally transforming to "data as a product". This is confirmed as well by [Ylijoki & Porras, 2016], stressing on that big data will have disruptive effects on firms, ecosystems and businesses, leading to emergence of new business models and value-creation mechanisms. The accumulation and integration of big data, coming from different sources is seen as a must for developing adaptive, smart, customer-oriented business models and processes. Therefore further approaches have to be developed enabling companies to integrate both data from internal processes as ERP/CRM/SCM systems, IoT sources and private clouds with industry data, social media, open government data and open public data, scientific data, partner's/ suppliers/competitor's/ecosystem data and others. Thus, the role of big data and data analysis for companies changed substantially during the stages of digital transformation, leading to further impact on industries and economy as a whole.

Making an overview of the successful digital transformation processes of old and big companies, Sebastian et al., (2017) identified three main elements. The first one is the adoption of digital strategy and technology-inspired value proposition. The main digital strategies can be either customer engagement (adopting technologies for increasing loyalty and trust) or development of digitized solutions (integrating combination of new products, services and data). The second element is to develop an operational backbone and the third element is to create digital service platform, enabling rapid innovation and responsiveness to new market opportunities.

In conclusion, the digital transformation is a complex issue, consisting of different aspects such as adoption of new digital strategy (digital transformation vision), framework (operational backbone), ecosystem approach, adoption of open environment for experiments.

3. Data value chain, Business models, Data ecosystem

The main data value-driven models are based on business analytics or data analytics methodologies, aimed to analyze, predict and control processes in business and industry [Coleman, 2016]. The main three subcategories of business analytics are: descriptive analytics (summarize, condense and aggregate data from complex data sets, using graphics and aggregated statistical metrics); predictive analytics (enable forecasts of future effects based on historical data, comprising statistical learning, machine learning, data mining and knowledge discovery from databases), and prescriptive analytics (transforming the results of descriptive analytics and predictive analytics into business decisions, based on optimization theory and operations research and quantitative tools) [Coleman, 2016]. It should be pointed out as well that a substantial pre-condition for any data based models and analytics is the good data quality [Baesens et al., 2014]. For these reasons, the data-value chain should include careful procedures and oversights to ensure high data quality through all data steps: (1) initial collection, (2) storage and updating, (3) retrieval, and (4) processing and preparation for analysis.

Investigating both research and practitioners' literature on open data business models, Zeleti et al., (2016) identified 15 business models: Premium, Freemium, Open Source, Infrastructural Razor and Blades, Demand-Oriented Platform, Supply-Oriented Platform, Free as Branded, Advertising, White-Label Development Cost Avoidance, Sponsorship, Dual Licensing, Support and Services, Charging for Changes, Increasing Quality through Participation, and Supporting Primary Business. In summary, Zeleti et al., (2016) conclude that the main open data business models are Freemium, Premium, Cost Saving, Indirect Benefit and Parts of Tools. The main value proposition out of the open data can be Usefulness, Process Improvement, Performance and Customer Loyalty. Roman et al. (2017) analyze that databased business models depend on data suppliers, (from simple data supplier to service provider), data sources (whether the data is internal or external) and what is done with the data (ranging from providing data for reuse, analyzing and aggregating existing data or even providing services). The revenue model for the different strategies are mostly subscription based but can also entail "freemium", pay-per-use or advertising. The more complex business models are also harder to execute, prompting open collaboration and co-creation [Roman, Liu, Nyberg, 2017]. Further, the authors determine that when an individual or organization is comfortable with data commercialization models they start to release more data and try to move on to more complex business models. Toots et al., (2017) find out that open data-driven service creation should be a process of value co-production, invoking collaboration between different stakeholders such as public administration, citizens and businesses. The authors propose to use open data for the co-production of new public services, or services leading to new pubic value, and their framework relies on the use of agile development practices in the creation of data-driven services.

The European data portal (EDP) Report 1 (2016) defines the following data-value chain processes and the main types of business users as follows:

- Data creation/Enablers: organizations, facilitating the supply or use of Open Data;
- Data aggregation/Aggregators: organizations that collect and aggregate open and proprietary data;
- Data analysis/Developers: organizations that design, build and sell web-mobile-apps based on data analysis;
- Data-based products and services/Data enricher: organizations that use Open data to enhance their products and services.

In the report of EDP (2016) among EU countries the data enrichers are among the most popular open data re-users (49,73%), and most of them work to improve company performance. Data aggregators (29,29%) works in new digital business blocks, and developers (18, 26%) are focused in customer touch points and new digital businesses. In the report from 2017, the EDP specifies further business cases for open data use and reuse and identifies more specific business models among open data end users. All four types of business models exploit open data in different ways to create value, for themselves, for their clients or for society. Thus, according to the EDP report (2017), the main revenue sources for OD businesses are: selling services (42%), selling products and services (21%) and selling products (10%). Most of the organizations, exploiting open data as resource are start-ups, and most of them are active in the information and communication sector. In this report, businesses can specify their activity as 1) Enhancing products, 2) Enhancing services, 3) Process optimization, 4) Data as a service, 5) Information as a service, 6) Answers as a service, 7) Development of web or mobile applications; or several of these. Thus the most popular Open data use is for internal process optimization, on the second place come organizations who facilitate access to and services on (aggregated) open data for others, on third place are organizations who offer (data analytical) products and services based on Open Data and finally come organizations who do not add financial value but create societal value. The main revenue streams are mainly coming from subscription fees, advertising, licensing, and consulting fees, lead generations and analytics fees (EDP, 2017). Confirming that, the statistics from 2016 (EDP, 2016) shows that open data in EU contribute mainly to performance management (internal, cost-savings), new digital businesses, developing new customer touch-points, improve customer understanding, company process digitization, top-line growth, digitally modified business and others. Open data re-use mainly helps companies to understand better the customer experience: for example, adapting marketing campaigns by geographic regions, making specific customer segmentations according to demographics, improve customer self-services and digital touch points and others. As most public administrations in general generate data about the environment, the legal system and the public safety, yet most of the data re-users are interested in the following sectors: government & public sector, economy & finance, transport and more particularly business registries and company data. The organizations re-using the Open Data work mainly in the IT sector, the public sector, financial and insurance services, as well in health, education & research, transportation, energy & utilities, culture & tourism and real estates.

3.1. Open data ecosystem

Open data can be provided both from public authorities and government structures or by third-level providers (as for example open scientific data, industry-specific data and others). Therefore, open data ecosystem can include both public and private providers, end-users and intermediary organizations. Based on the Open Data Consumer's Checklist (Open Data Institute), a model of value-adding activities of both open data publisher and open data intermediary is proposed and presented on table 1.

Table 1: *The role of the Data provider and Data intermediary for Open data value creation.*

	Data Provider	Data Intermediary
Access	Provide access Ensure data availability	Data format/enrich data use by API/
Ownership and licenses	Guarantee data origin Anonymize private data Appropriate license	Appropriate licenses for data re-use
Form	Provide data in machine-reading form, Quality meta-data Initial data processing	Summary, aggregated form, syntactic and semantic transformations; compatibility with the other sets
Quality	Current and frequently updated datasets; Long-term commitment	Ensure data accuracy Enrich data context Handle/recover missing data
Support	Data set documentation Meta-data analysis Support Reporting for errors	Supporting end-user Ensure feed-back; Good dataset documentation

Open data ecosystem can include different stakeholders, such as business organizations, public organizations, start-ups, universities, open spaces, public innovation labs, living labs, third level partners and others. In the whitepaper of Fraunhofer institute, Otto et al, (2016) highlight the role of Industrial data spaces that combine both industry (private) data, club data, and open data and promote the development of public connector that will orchestrate the data flows. The role of the public connector represents a third-level provider who can become as well a boundary organization that supports university-industry partnerships, involve various stakeholders, students, NGOs and public society organizations, organizes and hosts various events such as hackathons and competitions. In order that an Open data ecosystem is sustainable, the main success factors are publisher' sustainability, governance, financing models, appropriate technical architecture and the use of metrics.

3.2. Open Data technical requirements and formats

The Open data technical requirements and standards aim to ensure its further re-use in machine-readable format. Open datasets can differ between: low frequency use datasets (<5%), mid frequency use datasets (5%-10%), high frequency use datasets (>10%).The high frequency domains are in line with the high priority domains identified by the European Commission: such as geospatial data, earth observation and environmental data, transport data, statistical data and selected company data. Some other technical data can be identified as well:

- Common open datasets structures: tabular, hierarchical and network data structures.
- Data types can vary, including specific examples of data types such as legislative, statistical and geographic data that can require special treatment.
- Common Open Data formats and standards are CSV, JSON, GeoJSON, KML, XML and RDF Turtle.
- The most popular machine-readable data formats are: CSV, XLM or XLS;
- Interoperability of Open Data portals and metadata is ensured by standards such as DCAT-AP.

3.3. Barriers for Open Data implementation

Among the main barriers, still hindering wider Open Data re-use as driver of Digital Transformation can be identified [Coleman, 2016, EDP, 2016]:

- Lack of awareness for OD initiatives;
- Lack of knowledge and skills to use OD;
- No clear governance and lack of responsible person within an organization, as a company data manager;
- Data quality: lack of capacity to combine open data with internal data;
- Lack of capacity for internal data management;
- Lack of appropriate data license.
- Big data architectures clearly represents an apparent barrier for SMEs, both from a financial and cultural point of view [Coleman, 2016].

4. Open data in Bulgaria

Bulgaria ranks among the EU trend setters for its Open data adoption, as identified in Open data maturity report (EDP, 2017). The open data portal in Bulgaria is launched in November 2014

following the Directive 2013/37/EU of the European Parliament and European Council, amending Directive 2003/98/EC on the re-use of public sector information text with EEA relevance. The adoption of the new amendments of the EU Directive induced the amendments in the Bulgarian Law for public information access, along with other changes in the legal framework. With decision of the Council of Ministers in 2015 is approved a list with 119 datasets, and additional 149 data sets have to be made publicly available till the end of 2017. Currently 6983 datasets are available on the Bulgarian OD portal https://opendata.government.bg/ from 490 registered public data providers.

The Open data initiative in Bulgaria started in 2014 by a working group under the Council of Ministers' administration [Gerunov, 2015]. On Bulgarian OD portal, datasets are freely available for commercial and non-commercial use. The Open data project is developed on the open-source CKAN platform (https://ckan.org/), supported by Open Knowledge Foundation. Citizens and data-users have additional possibility to send feedback, to request public data, to specify public data that should be available in machine-readable data and other.

Among the main barriers for Open Data re-use in Bulgaria, the EDPa (2017) report identified two main issues: the lack of awareness and technical barriers, such as the low quality of data sets and low synchronization of information in the databases. Technical obstacles exist to automatically upload and update data with administrations that maintain and collect the information. Data users in Bulgaria still perceive the quality of Open Data to be low [EDPa, 2017]. The low quality refers to both the data itself as well as the accompanying metadata, and the lack of standardization users to develop permanent solutions to re-use Open Data in their processes.

Conclusions

The present research provides an overview of the main mechanisms for open data value creation. Among the main recommendations for companies are to start their digital transformation by developing open data projects, focused on performance management, developing customer touch points or starting new databased digital business. As the availability of open data will increase both in Bulgaria, in EU and internationally, defining now a company open data value strategy can become a trigger for further digital transformation. Open data strategies can help companies to improve customer experiences, to facilitate adoption of Cyber-physical systems and to enhance data-driven decisions, providing complex models for data re-use from different sectors and application areas.

Acknowledgement

The authors would like to acknowledge the support by the Interreg DIGITRANS: Digital transformation in the Danube region - http://www.interreg-danube.eu/approved-projects/digitrans.

References

Almada-Lobo, F. (2016). The Industry 4.0 revolution and the future of manufacturing execution systems (MES). Journal of Innovation Management, 3(4), 16-21.

Baesens, B., Bapna, R., Marsden, J. R., Vanthienen, J., & Zhao, J. L. (2014). Transformational issues of big data and analytics in networked business. *MIS quarterly*, *38*(2), 629-631.

Carrara, W., Chan, W. S., Fische, S., & Steenbergen, E. V. (2015). Creating Value through Open Data: Study on the Impact of Re-use of Public Data Resources. European Commission.

Country Factsheet for Bulgaria, European Data Portal, https://www.europeandataportal.eu/sites/default/files/country-factsheet_bulgaria.pdf

Davies, T., & Perini, F. (2016). Researching the emerging impacts of open data: revisiting the ODDC conceptual framework. The Journal of Community Informatics, 12(2).

Dawes, S. S., Vidiasova, L., & Parkhimovich, O. (2016). Planning and designing open government data programs: An ecosystem approach. *Government Information Quarterly*, *33*(1), 15-27.

Digitizing European Industry EC portal, https://ec.europa.eu/digital-single-market/en/policies/digitising-european-industry

EDP, 2016, Analytical Report 1: Digital Transformation and Open Data, available on https://www.europeandataportal.eu

EDP, 2016, Open Data Goldbook for Data Holders and Data Managers, available on https://www.europeandataportal.eu/sites/default/files/goldbook.pdf

EDP, 2017, Open Data Maturity in Europe, available on https://www.capgemini.com/wp-content/uploads/2017/11/report-open-data-maturity-in-europe-2017.pdf

EDPa, 2017, Analytical Report 5: Open Data Barriers, available on https://www.europeandataportal.eu

European Commission. 2011. "Open Data: An Engine for Innovation, Growth and Transparent Governance." Brussels. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2011:0882:FIN:EN:PDF.

Gerunov, A., 2015, Open Data: Policy and Implementation in Bulgaria. *Big Data, Knowledge and Control Systems Engineering*, 1-11.

Huijboom, N., & Van den Broek, T. (2011). Open data: an international comparison of strategies. *European journal of ePractice*, *12*(1), 4-16.

Open Data Handbook, Open Knowledge International, 2015, http://opendatahandbook.org/en/what-is-open-data/index.html

Open Data Institute, Open Data Consumer Checklist, https://theodi.org/guides/the-open-data-consumers-checklist

Otto, B., Juerjens, J., Schon J., Auer S., Menz N., Wenzel S., Cirullis J., (2016), Industrial Data Spaces, Fraunhofer Institute whitepaper available on: https://www.fraunhofer.de/content/dam/zv/en/fields-of-research/industrial-data-space/whitepaper-industrial-data-space-eng.pdf

Roman, M., Liu, J., & Nyberg, T. UNIVERSITY-INDUSTRY COLLABORATION THROUGH BIG DATA AND OPEN SCIENCE. ICEIRD conference proceedings, *31 Aug–1 Sept 2017*, 371.

Saldivar, A. A. F., Li, Y., Chen, W. N., Zhan, Z. H., Zhang, J., & Chen, L. Y. (2015, September). Industry 4.0 with cyber-physical integration: A design and manufacture perspective. In Automation and computing (icac), 2015 21st international conference on (pp. 1-6). IEEE.

Sebastian, I. M., Ross, J. W., Beath, C., Mocker, M., Moloney, K. G., & Fonstad, N. O. (2017). How Big Old Companies Navigate Digital Transformation. MIS Quarterly Executive.

Sommer, L. (2015). Industrial revolution-industry 4.0: Are German manufacturing SMEs the first victims of this revolution? Journal of Industrial Engineering and Management, 8(5), 1512.

Toots, M., McBride, K., Kalvet, T., Krimmer, R., Tambouris, E., Panopoulou, E., ... & Tarabanis, K. (2017, September). A Framework for Data-Driven Public Service Co-Production. In *International Conference on Electronic Government* (pp. 264-275). Springer, Cham.

Welle Donker, F., & van Loenen, B. (2017). How to assess the success of the open data ecosystem? *International Journal of Digital Earth*, *10*(3), 284-306.

Ylijoki, O., & Porras, J. (2016). Perspectives to definition of big data: a mapping study and discussion. Journal of Innovation Management, 4(1), 69-91.

Zeleti, F. A., Ojo, A., & Curry, E. (2016). Exploring the economic value of open government data. *Government Information Quarterly*, *33*(3), 535-551.