APPLICATION OF OPEN DATA IN THE CZECH REPUBLIC

[a]LUCIE STAŇKOVÁ, [b]JAN BOUBÍN

Technical University of Liberec, Studentská 1402/2, 461 17 Liberec 1, Czech republic
email: [a]lucie.stankova1@tul.cz, [b]jan.boubin@tul.cz

The article was developed with the support of SGS project of Technical University of Liberec No. 3600/21194 "The use of Open Data for the purpose of implementing and managing Start – up projects and increasing corporate transparency".

Abstract: Open data represent a significant potential for a business development through strategic planning and innovation management. They can also serve as a tool for increasing transparency and reducing corruption. Yet, open data are still neglected and they are not given enough importance. This might be caused by insufficient information support of open data. The paper aims to analyze the current level of application of open data by a primary questionnaire survey on a sample of private sector companies and public sector organizations in the Czech Republic. The results of this paper represent a comprehensive overview of current open data applications by companies and organizations in the Czech Republic.

Keywords: application, Czech Republic, open data, primary survey, private sector, public sector

1 Introduction

Open data brings new potential for economic development. Wide range of data usage include, for example, strategic planning, innovation management, or corporate transparency. Open data may be important for research and development within businesses and across sectors. Last but not least, they can be an important source of initial information for a new business or project (for example, a start-up).

The digitization of the current market is evident from a number of initiatives. The most significant is Industry 4.0 which considers digital data sharing to be very important. Open data support and emphasizing its importance should in the future lead to streamlining business processes and also to creating more diverse partnerships, alliances or clusters. In the future, open data could, for example, help consumers with choosing the right product from a trustworthy company (with emphasis on its good reputation).

Different experience in working with open data is due to the different nature of subjects. In the public sector, data are published in a variety of forms and in the long-term period (which is determined by law). The work with them becomes an integral part of public organizations functioning. In the private sector, open data is a new tool, and therefore, the awareness of it is not at the level where it could be.

Some companies do not have enough information on the data-opening option, other companies cannot efficiently work with data and use their potential in a day-to-day process practically. Apparently, information support for open data in the private sector is not at a very high level. Current initiatives target mainly at the public sector, especially at governmental institutions. Various organizations often emphasize the positive effects of open data across sectors and market segments. However, there are also shortcomings in the use of open data in the public sector.

2 Problem formulation

The basic characteristics of open data is their availability. The data is freely accessible electronically to all businesses, institutions, and private entities. This enables their free distribution and the possibility of sharing with as many entities as possible. Specifically, these are data from a variety of areas of economics, economy, research, development, etc. (Open Knowledge Foundation, 2012)

In the Open Knowledge Foundation handbook (2012) it is stated that availability, redistribution, re-use, a lack of technology constraints, authorship and integrity are one of the basic prerequisites for open data. It also states that the prohibition of discrimination against persons, groups or open data applications must be respected. The rules also apply to license requirements. The license must not restrict the dissemination of further work (or the same license must be used for the new document as for the original document). The authors of the handbook also claim that the rights associated with a particular dataset must apply to all users without the need for an additional license.

Berg (2013) claims that open data can be used by a variety of entities. The users of the open data are divided into five basic groups - open data suppliers, data aggregators, application developers, enrichers and enablers.

Suppliers publish open data freely. It may not be their primary goal; however, it is often a form of a company strategy. Data aggregators mainly process external data. Their main benefit is the added value of working with these data in the form of their further processing, cleaning or combining. Application developers are a very important component because they are the ones who provide data processing and make open data accessible. The primary goal of developers is to process data into usable applications as efficiently and organized as possible. Data enrichers collect and then create their own expertise. This creates added value for customers, making primary and austere open data more user-friendly. Enablers create tools, technologies or methods by which data are published and can become open. The supplier, the aggregator and the enricher can be identified in the private and public sectors. Application development and data access are typically provided by private companies (mostly in the field of information technology). (Berg, 2013)

Open data are widely used in open innovation. Open innovations represent a situation where a company is willing to share its knowledge and experience to some extent with other entities and continue to develop the innovation potential. In the open innovation model, shared data are crucial. Based on open data, experts can deduct or calculate the success or return of the planned innovation. They can, therefore, come to a conclusion on whether it is cost-effective to introduce the innovation. (Boček, 2012, Chlapek et al., 2012)

Open data have a significant impact on the development of the economic environment of a country as well as on individual companies. Open data provided by external entities may lead to build an optimal company strategy, to change strategies or to search for new business activities and opportunities. (Lathrop and Ruma, 2010; MV ČR, 2015) This supports the innovative thinking of corporate employees. Through open data, the suitability of the innovation and further the innovation itself can be adequately planned and managed. (Saebi and Foss, 2015) Furthermore, open data can serve as a tool for evaluating and searching for potential partners, increasing corporate transparency, or communicating with customers. (Open Knowledge International, 2016)

According to Chlapek et al. (2012), the most important objective of the public sector is to increase the transparency of activities of public organizations. Strengthening the reputation of public organizations is closely linked with the funding of these entities, which is ensured primarily through collected taxes. That is why the clarity and awareness of a number of orders is very important and increases the credibility of the country and public organizations.

3 Methods

The main research question is: To what extent and in what areas are open data applied by companies and organizations of the private and public sector in the Czech Republic?

The aim of the paper is to evaluate the current level of open data application in the Czech Republic with an emphasis on their use

in planning, management and decision making of companies and organizations of a private and public sector. The paper consists of the research of available sources on the given issue, the questionnaire primary survey and the analysis of the data obtained from the primary survey including the use of statistical methods.

Within the project, a primary research was conducted on the use of open data by the private and public sector in the Czech Republic. For this purpose, the questionnaire, which contained 15 questions, was created. The first part of the questionnaire was focused on identifying companies and organizations. Other questions divided respondents into those who know open data and those who had not worked with them yet. In conclusion, it was investigated for what purpose and what data companies and organizations use, provide and miss.

Before the survey itself and sending the questionnaire, a pilot study was conducted. This study was designed to validate and improve the questionnaire. Five companies were asked to complete the questionnaire and send comments. The observations received were incorporated into the questionnaire.

4 Problem analysis

The questionnaire was sent to 3,650 e-mail addresses, of which 981 were undelivered. For companies, a multi-step selection was made. It was determined how many companies were located in what regions and by this ratio the companies were randomly selected by the MagnusWeb application. This is the application in which data on business entities of the Czech Republic are collected. As the questionnaire contained a question of the size of the company in terms of number of employees, the same number of companies from each category was included in the survey, that is, the number of employees up to 50, the number of employees from 50 to 250 and the number of employees over 250. Altogether 3,000 companies were addressed.

Public sector organizations were selected according to the type of an entity. Ministries, public universities, regions and central state administration bodies were included in the basic set, so all these institutions were addressed. Other state administration bodies, interest groups, civic associations, unions, churches, foundations, charitable organizations and allowance organizations were selected at random. From each category there was addressed up to 10% of organizations. A total of 650 institutions from the public sector was included in the survey. The return of the questionnaire was a total of only 3.8%, i.e. 140 completed questionnaires. There were received 73 responses from the public sector, which is more than 11% return. Thus, the public sector survey output can certainly be considered valid. However, only 2.2% of completed questionnaires were sent back by private sector companies. It is, therefore, necessary to take the corporate sector with caution as with more questionnaire replies, the results could vary.

Table 1 provides a brief overview of the structure of companies and organizations responding to the questionnaire. The responses of the public sector with the highest return from the municipalities, regions and state organizational units prevailed. As far as the rest of the responses is concerned, the municipality was included once, the respondent apparently overlooked that this option was mentioned in the selection, one response came from the legal entity and one from the representative of the library. The private sector divided by the number of employees was represented by approximately the same proportion. The organizations such as other services, public administration, agriculture and government and self-government were most willing to fill in a questionnaire. In Table 1 only those other replies that were answered more than twice are given. Other respondents included fields such as regional development, IT, consultancy, sports and physical education, waterworks engineering or pharmacy.

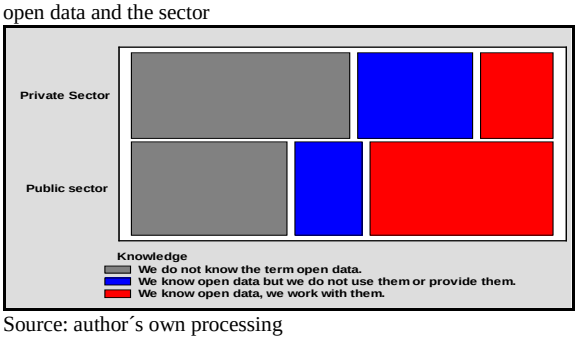
Table 1: Comprehensive information on the number of replies of addressed businesses and organizations

Company/ organization	Division of companies/ organizations	Company/ organization field
Private sector: 67 (47.9 %)	No. of employees up to 50: 23 (34.3 %) No. of employees from 50 to 250: 27 (40.3 %) No. of employees over 250: 17 (25.4 %)	safety: 1 (0.7 %) tourism/leisure time: 1 (0.7 %) transportation: 2 (1.4 %) ecology: 1 (0.7 %) financial services: 9 (6.4 %) culture: 4 (2.9 %) other services: 17 (12.1 %) food industry: 2 (1.4 %) craft activities: 2 (1.4 %) social services: 9 (6.4 %) construction: 4 (2.9 %) technical services: 1 (0.7 %) education: 3 (2.1 %) healthcare: 3 (2.1 %) agriculture: 11 (7.9 %) manufacturing services: 2 (1.4 %) others: 68 (48.6 %) public administration: 17 government and self-government: 10 engineering: 5 electrotechnical industry: 4
Public sector: 73 (52.1 %)	civic association: 1 (1.4 %) registered church/religious group: 2 (2.7 %) foundation: 2 (2.7 %) charitable organization: 1 (1.4 %) public university: 1 (1.4 %) territorial self-governing unit (municipality/region): 37 (50.7 %) state organizational unit: 23 (31.5 %) allowance organizations: 3 (4.1 %) others: 3 (4.1 %): municipality: 1 legal entity: 1 library: 1	

Source: author´s own processing

After the distribution of the respondents by sector, size, types and fields, the respondents were further divided according to their experience with open data. It appeared that 45.7% of the respondents had not met the concept of open data yet. The term of open data was familiar to 54.3% of the participants. However, only 32.1% actively worked with open data. The questionnaire revealed that open data were more familiar and useable in the public sector. Out of 45 responses of the respondents who regularly work with open data, there were only 12 companies. This assertion was confirmed by the test of the independence of variables between the knowledge of open data and the business or organization sector, whether it was the private or the public sector. In Table 2, Chi-square test values of the test are summarized which confirm that the variables are not independent. The public sector is more aware of open data, which is also noticeable from the graph shown in Figure 1.

Figure 1: Graph of independence test between experience with open data and the sector



Source: author´s own processing

Table 2: Test of independence of variables – experience with open data vs. sector/field

Test	Statistical value	Degrees of freedom	P-Value
Chi-square	12.146	2	0.0023

Source: author´s own processing

Attention was also paid to the independence between the experience with open data and a company size. Again, a Chi-square test was used and it confirmed that the variables are independent. The values are shown in Table 3. Businesses of all sizes work with open data.

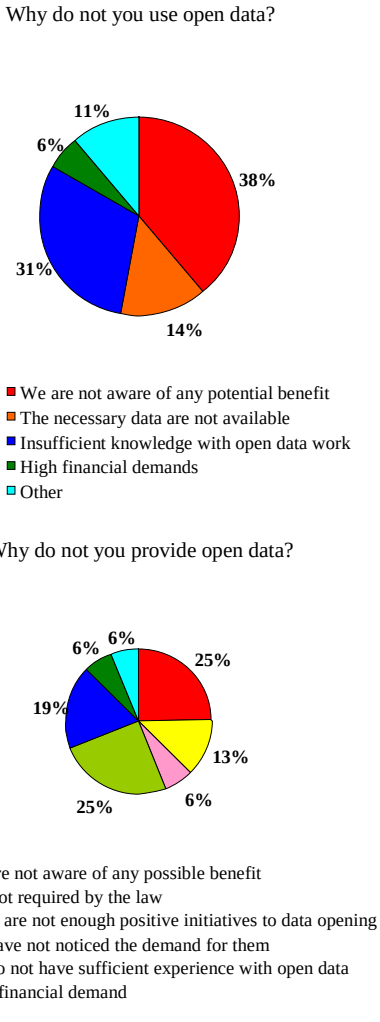
Table 3: Test of independence of variables – experience with open data vs. company size according to no. of employees

Test	Statistical value	Degrees of freedom	P-Value
Chi-square	1.614	4	0.8063

Source: author´s own processing

Figure 2 gives an answer to the question why the companies and the organizations do not use and do not provide open data. This question was given to such respondents who know open data but do not use it or provide it. The most common answer was that the companies were not aware of any potential benefits of open data. So, it can be assumed that it is a new tool that needs to get into the subconsciousness of companies and organizations first. Another very common answer was that the respondents did not have enough experience with open data. The public sector, which is more experienced in this issue, could provide better information about the availability of open data via media, the Internet or applications resulting from the use of open data.

Figure 2: Reason for not using and not providing open data

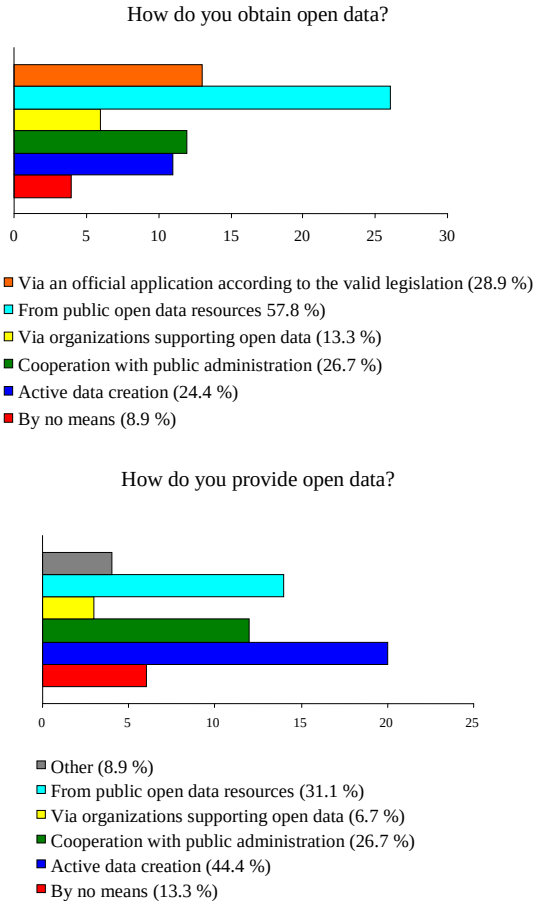


Source: author´s own processing

With the help of another question the respondents who work with open data were divided into different user groups. Some of the respondents deal with more than one open data work. Out of 45 regular open data users, most companies and organizations provide data. This option was selected by 34 respondents, i.e. 75.6%. The second most common activity with open data is to make them accessible. 31.1% of the respondents, namely 14 companies and organizations, deal with the creation of tools that allow data to be opened and made accessible. Data aggregators were represented by 28.9%. Thirteen respondents deal with data collection and processing, followed by the creation of added value. Eight information enrichers who focus on making expert opinion to gained data also participated in the survey (17.8%). Application developers had the weakest representation in the survey only by 11.1%, that means four companies and one public sector organization.

The survey also focused on how companies and organizations gained and provided data. As can be seen from Figure 3, most respondents receive data from public sources, official requests under valid legislation, cooperation with public administration or active data generation. In the case of data provision, activities such as active data generation, provision from public sources or cooperation with public administration predominate.

Figure 3: How to gain and provide open data



Source: author´s own processing

Further, it was explored from which areas the companies and the organizations use open data and make them accessible. Table 4 summarizes the values obtained from the survey. The companies and the organizations mostly use data from statistics, public procurement, maps, business registers or legislation. Data providers, who only make data available and do not use it, selected the answer ‘other’. Another interesting answer was that thanks to the wide range of services offered, the company used almost all open data. The open data provided by respondents include budget, statistics, public procurement and election results. Also, in the case of making data accessible, some respondents chose the answer ‘other’ because they did not

make data accessible, but only used it. Other responses also included contracts, employment, social services and security or investment. One of the respondents stated that they only made accessible those data that were legally required by law.

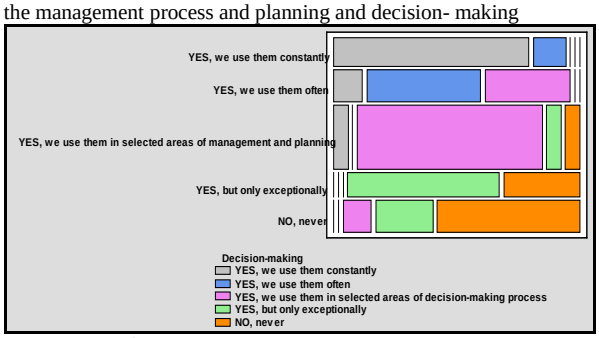
Table 4: Areas of data use and data accession

Data areas	Data use (no. of replies)	Data accession (no. of replies)
Maps	19	10
Land	4	1
Statistics	24	19
State budget/budget	13	22
Government expenditures	9	0
Business register	18	2
Legislation	15	5
Public transportation	8	2
Business	3	1
Healthcare	2	0
Education	9	4
Crime	2	2
Environment	13	4
Election results	11	11
Public procurement	19	19
Others	9	17

Source: author´s own processing

Applying open data in the process of management, planning, and decision making is fairly common among active users of open data. 33.4% of the respondents applied open data in the management and planning process continuously or very frequently. In the decision-making process, representation was slightly lower, 28.9%. In selected areas of management and planning, open data were used by 35.6% of the companies and the organizations. Only 17.8% of the respondents never used open data in the management and planning process. Others, who know open data, only used them in exceptional cases. In selected decision-making processes, 37.8% of the respondents used open data. 17.8% of the decision-makers also stated that open data never helped them. Only in exceptional cases, 15.5% of the companies and the organizations used open data in the decision-making process. Using the Chi-Square Test of Independence of Variables, it was proven that the companies that used open data in the management and planning process usually did so for decision-making (statistical value – 77.976, P-Value – 0.0000). The conclusions are shown in Figure 4.

Figure 4: Proof of dependence between the use of open data in the management process and planning and decision- making



Source: author´s own processing

Three open questions were used in the questionnaire. For example, in response to the purpose/method of using open data, these responses were as follows:

- for the need of self-government and state administration,
- in the performance of public administration,
- for the company´s own and public information,
- for education,
- in the decision-making process,
- transparency.

It is clear from the answers that the companies and the organizations use data primarily to increase transparency and awareness. They do so to improve processes within the company and to share data with the public more effectively. However, there were also replies that the respondents do so for legislative reasons. These entities publish the data only because the law imposes it and because they had not revealed their benefits yet.

Another open question confirms the conclusions of the previous question. Respondents gave the answer to the question of the open data benefits. The companies and the organizations see the biggest benefit in transparency and awareness. Open data provide quick and almost immediate information. Another repeated response was to simplify the processing of requests for information under the current legislation. This fact shows that they publish the data because they are required to do so by one of the laws. A very frequent answer was that the companies and the organizations see no benefit in open data. This response clearly prevailed in the private sector. This may be caused by insufficient experience with open data. On the other hand, a positive opinion that open data are crucial for some businesses and organizations prevailed in the public sector organizations.

The last of the open questions addressed data issues that the companies and the organizations do not use or cannot use. The most common answer was that the companies and the organizations did not know about such data or that such data did not exist. The fairly common answer was that all the data they needed and wanted to use were available to them. Other businesses and organizations would appreciate more accessible information on public and state administration, public procurement, timetables, legislation, hazardous waste, labour offices, education, and crime. Some companies would also like to use more data from their customers, suppliers, subscribers and competitors.

At the end of the questionnaire, a question was laid to evaluate the main obstacles in gaining and the use of open data. The answers are summarized in Table 5. It shows the number of answers and the percentage out of 76 respondents who know open data and chose the given answer. The biggest problem is that the data are not regularly updated. Another obstacle can be seen in public and online data availability. A positive fact is that the least number of responses were given to the question of whether data is not in the required quality and format. This suggests that the companies and the organizations learn to publish data in a better form than just scanned documents, as is often the case today. Among other responses, for example, there appeared the view that an obstacle to open data is the general awareness of the possibility of their use and public awareness. The problem is also the lack of opening data anchoring in the valid legislation. At present, opening data is mostly a voluntary matter. However, there were also respondents who did not know any obstacles in obtaining or using the necessary open data or they were not aware of it.

Table 5: Obstacles in gaining and using the necessary open data

Answers	No. of answers	% of answers by respondents who know open data
Data do not exist	11	14.5 %
Data are not in an electronic form	18	23.7 %
Data are not publicly accessible	27	35.5 %
Data are not available online	24	31.6 %
Data are not current and regularly updated	31	40.8 %
Data are not consistent	15	19.7 %
Data are not in a required quality	10	13.2 %
Data do not have	10	13.2 %

a required format		
Financially demanding	12	15.8 %
Other	12	15.8 %

Source: author´s own processing

The survey shows that open data is a tool that companies and organizations are currently learning to use. It seems that public sector organizations are one step ahead. However, even in the open data application, they still have to improve. In the field of open data, it would be advisable to get inspired by countries that are leaders in this branch. Great Britain and the US have been using this tool for quite some time and in a quality way. Companies and organizations in the Czech Republic could find the motivation there to use open data in their activities.

5 Conclusion

One of the major issues of primary research conducted through the questionnaire includes a low return on the answered questionnaires. Public sector organizations participated to a fairly large extent, the return on the questionnaire was over 11%. On the contrary, the companies in the private sector did not show much interest. In view of these results, it may be interesting and beneficial to investigate why companies do not engage in research and what would motivate them to participate in surveysmore often.

Open data can be used as a competitive advantage. Based on these data, acompany can control the processes inside its business, create and build its good name. In this way, the company has the opportunity to acquire new customers and build the loyalty of those existing. (Open Knowledge International, 2016) A wide range of data offers the potential for economic development of companies and the elimination of risks that may eventually lead to financial problems of the company or its extinction. Open data offers the possibility of reducing the risk, which should strengthen not only the development of the company but also the whole market structure. (Solom and Björk, 2012)

To find the level of open data application in the Czech Republic by the privateand the public sector, the primary research was conducted in the form of a questionnaire survey. There were addressed 3,000 private sector companies and 650 public sector institutions. The main issue of the research conducted by email questionnaire is its low return. This was the case with this research, where the return was only 3.8%. For the public sector, results can be considered valid, as returns were over 11%. However, the private companies returned only 2.2% of the questionnaires. Conclusions made for the private business sector need to be treated with caution.

The research results show that open data are now more widely applied by the public sector. This is probably due to the little experience with open data that private businesses have.

Statistical testing of the obtained results demonstrated that companies of all sizes work with open data. This tool is applicable to small, medium and large enterprises. In addition, it was found that when the companies and the organizations use open data in the management and planning process they do it mostly in a decision-making process. This demonstrates the complexity of using open data.

The research has also shown that there is insufficient awareness of open data. The most common answer to the question why companies and organizations did not use data and did not provide them was that they saw no potential benefit in open data. It can therefore be assumed that without information support the development of open data will be complicated. The benefits of open data are often neglected. Increased awareness of open data could, therefore, lead to an increase in entities that would use open data. This is followed by the anticipated faster growth of the economy. Therefore, greater awareness of open data is needed, with an emphasis on positive impacts on market development and also companies and institutions themselves.

In the Czech Republic, open data are at the beginning of development and their application is not sufficient according to research results. There is a lack of support from the Czech Republic and the European Union. This may result in a slower economic growth. Nor is achieved the innovation potential of the economy. Data sharing is perceived rather negatively, data potential for future growth is overlooked.

On the whole, open data should be more prominent across companies and organizations of all sectors to enable effective economic growth and increased transparency of economic processes. Public sector organizations are legally required to open data. For private sector companies, this aspect is missing, with the exception of the disclosure requirement for some industries and the financial statements for all companies. The incentive for data opening could, therefore, be their legislative support with an emphasis on the development of the business environment.

Literature:

1. Berg, M.: *Otevřená data a jejich byznys modely: kde v nich hledat peníze* [online] [cit. 2016-08-16]. 2013. Available from: <http://www.datablog.cz/clanky/byznys-modely>.
2. Boček, J.: *Otevřená data ve státní správě: Nová éra rozhodování* [online] [cit. 2016-05-18]. 2012. Available from: <http://www.osf.cz/wp-content/uploads/2015/08/ODSS_Text_web_01.pdf>.
3. Chlapek, D., et al.: *Metodika publikace otevřených dat veřejné správy ČR* [online] [cit. 2016-10-09]. Prague, 2012. Available from: <http://www.otevrenadata.cz/res/data/002/003544.pdf>.
4. Lathrop, D., Ruma, L.: *Open Government: Collaboration, Transparency, and Participation in Practice*. Newton: O'Reilly Media, 2010. 432 p. ISBN 978-0-596-80435-0.
5. MV ČR: *Otevřená data* [online] [cit. 2016-06-14]. Prague: Ministry of the Inerior of the Czech Republic, 2015. Available from: <http://www.mvcr.cz/clanek/otevrena-data.aspx>.
6. Open Knowledge Foundation: *Open Data Handbook* [online] [cit. 2016-06-03]. United Kingdom: Open Knowledge Foundation, 2012. Available from: <http://opendatahandbook.org/en/>
7. Open Knowledge International: *Vision and Values* [online] [cit. 2016-06-23]. 2016. Available from: <https://okfn.org/about/vision-and-values>.
8. Saebi, T., Foss, N.: *Business models for open innovation: Matching heterogeneous open innovation strategies with business model dimensions.* vol 33, no. 3. European Management Journal, 2015. pp. 201-213. ISSN 0263-2373.
9. Solom, D., Björk, B.: *A Study of Open Access Journals Using Article Processing Charges.* vol. 63, no. 8. Journal of American Society for Information Science and Technology, 2012. ISSN 2330-1643.

Primary Paper Section: A

Secondary Paper Section: AE, AF