



The 7th International Conference on Ambient Systems, Networks and Technologies  
(ANT 2016)

## Agile project management approach and its use in big data management

Patrícia Franková<sup>a</sup>, Martina Drahošová<sup>b</sup>, Peter Balco<sup>a,b\*</sup>

<sup>a</sup>Comenius University, Faculty of Management, Odbojárov 10, P. O. Box 95, Bratislava, 82005, Slovakia

<sup>b</sup>Comenius University, Faculty of Management, Odbojárov 10, P. O. Box 95, Bratislava, 82005, Slovakia

---

### Abstract

The use of project management methodology for each project generates a lot of questions as well as recommendations. One group argues that the standard methodology are applicable to all projects. Another group prefers the view that reflects the project management methodology tailored to the project. In our contribution we summarize the practical views of project managers on project management issues in the area of big data management.

© 2016 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Peer-review under responsibility of the Conference Program Chairs

*Keywords:* Agile big data management; project management; big data projects; practical recommendations.

---

### 1. Introduction

The aim of this work is to document agile approach to project management and suggests ways of using it in projects related to Big Data management. For this purpose we conducted interviews with experts in the field of big data management and project management, which was held in order to clarify the issue of Big Data management and the possibility of applying agile approach to big data projects in different companies. Based on the evaluation and processing of data from the interview, we identify which of the principles of Agile Manifesto, can be used in the management of Big Data projects. Subsequently we determined the recommended practices in the management of Big Data and Big Data projects management. We based our analysis of the opinions of practitioners and the analysis

---

\* Corresponding author. Tel.: +421 915 995 755  
E-mail address: [peter.balco@fm.uniba.sk](mailto:peter.balco@fm.uniba.sk)

of the questionnaire. The questionnaire survey was designed to generate practical knowledge and experience in the field, which was subsequently distributed to companies undertaking the big data projects.

By analyzing the data obtained from interviews and questionnaire survey, we try to answer the following questions:

- What approach is suitable for the management of Big Data projects?
- Is it possible to use an agile approach to managing Big Data projects?
- What are the practical recommendations for the implementation of Big Data projects?

Project management can be understood as "the planning, delegation, monitoring and control of all aspects of the project and motivate the participants to achieve the project objectives within the expected performance parameters set for time, cost, quality, scope, benefits and risks".<sup>1,2</sup> Similarly, standard ISO 1006<sup>3,4</sup>, defines project management as "the planning, organizing, monitoring, managing and reporting of all aspects of the project and the motivation of all project participants to achieve the project objectives." In principle, it is the implementation and control of the actions that are necessary during project implementation in order to achieve the project objectives within pre-defined limits<sup>28,31</sup>.

## 2. Agile approach to software development

In addition to the plan managed approach of software development is gradually focus on an agile approach. Word agile has become one of the modern concepts of the 21st century. When we were interested in the popularity of agile approach and agile framework Scrum, we searched for these terms in the English language between trends in the search engine of company Google<sup>5</sup>.

What does the term "agile" mean? This term is not in the context of software development clearly defined. The origin of the term of agility in manufacturing is defined as "the ability to successfully sell low-cost, high-quality products with short delivery times and in different capacities that by the adaptation to the needs of the user provide increased value for customers".<sup>6</sup> In the article<sup>7</sup> it is stated that "a key precondition for agility in software development is the freedom to adapt procedures and methods to the needs of a particular project."

Success of practicing agile development is confirmed by the latest research State of Agile™ of<sup>8</sup> sponsored by Version One, which a supplier of business software in the field of agile is handling. Of the more than 3,500 respondents surveyed, mainly in the field of software development, up to 88% using agile development. An annual increase of 2 percent compared to the previous two years also reflects the growing popularity of agile approaches.

### 2.1. Agile Methods

Agile methods are "methods that are trying to focus on the primary objective of effective software development, ie. the creation of working software (without defects)".<sup>9</sup>

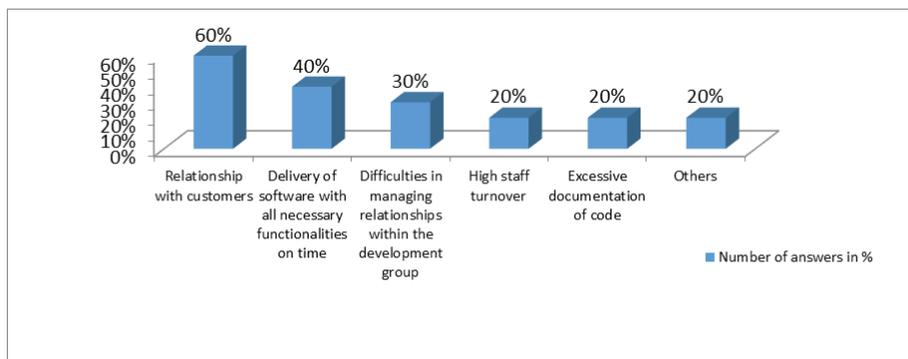


Fig. 1. The transition to agile methods during the software development mitigated these problems.

Effective software development means that we are developing the software "sufficiently" for a given situation, assuming that the requirements will vary<sup>26, 29</sup>. The term "sufficient" thus indicates that "the development is not necessary to do anything more than is agreed with the customer and sufficient for a given iteration".<sup>10</sup> According to research<sup>11</sup> was by the transition to agile methods during the software development mitigated the problems presented on the figure 1.

### 3. Big Data Management (BDM)

Big Data management can be understood as "the organization, administration and management of large volumes of structured and unstructured data".<sup>12</sup>

Research by Gartner confirmed the growing interest in Big Data and the increasing percentage of their deployment in the companies. In 2012, it invested or planned to invest in Big Data technology 58% of companies<sup>13</sup>, in 2013 64% and in 2014 even to 73% of surveyed organizations.<sup>14</sup> On the other hand, research has shown that in 2013 had deployed Big Data projects 8% and in 2014 only 13% of surveyed organizations. Although they experienced annual growth of the deployment of 5%, it is still a small amount compared to plans<sup>27, 30</sup>. The above suggests that Big Data offers to organizations a non-negligible value, but are also difficult to implement, not only in terms of technology and investment, but also need to change the strategic thinking of the whole company.

#### 3.1. The definition of the Big Data

The term Big Data can we can generally introduce as a "massive data files that have a large, diverse and complex structure and are therefore difficult to store, analyze and visualize for other processes and results".<sup>15</sup> It is not possible to clearly determine the size range for the definition of Big Data. Big include "data files with a size that exceeds the ability of commonly used software tools these data capture, manage and process in acceptable total time. Large size data are constantly moving target from a few dozen terabytes to many petabytes of data in one data set".<sup>16</sup>

The individual attributes characteristic of Big Data suggests the definition "Big Data are large-volume, high-speed and widely diverse information assets that require new forms of processing to allow improved decision-making, greater insight into the discovery and optimization of processes".<sup>17</sup> The first part of this formulation suggests that Big data is from the data of traditional analysis characterized by three main components: volume, velocity and variety. The analyst Doug Laney<sup>18</sup> in his work suggests this popular 3 dimensional framework for defining Big data, often called the "3VS" model. This name is derived from the first letter "V" of the components of the model (volume, velocity, variety). Let us then more detail describe what is meant by the components of the framework, which became a general test to determine whether it is appropriate to adopt a Big Data approach for data analysis.

### 4. Practical recommendations for the management of Big Data projects

In this section, we will build upon the resources that make up the knowledge of scientific articles, books and blogs, which is based on theoretical part. Furthermore, we build on the analysis of the contributions of the conference and analysis of Big Data management questionnaire from our research. We have submitted a questionnaire to participants of the conference and also send to people working in the field. A sample of 237 respondents represent people with an average 5-years of experience who work in organizations with different number of employees in positions such as: CEOs, managers at various levels, assistants and IT development projects.

The majority of respondents, up 40.5%, in the companies have established plan managed approach. In the agile controlled environment, working 24.3% and the rest of respondents could not answer this question or not selected any management approach.

Up to 2/3 of the respondents prefer the internal approach to Big Data implementation, 1/3 is considered preferable hosted platforms.

It is interesting that according to research of people working in IT field, it is also found that 66% of respondents prefer the internal approach of implementation of Big Data projects.<sup>24</sup> On the other hand, the same research

indicates that for 61% of respondents is agreed solution based on a hosted platform. The reason is to gain time to create business value.

The questions in the questionnaire were to focus on the priority issues of Big Data and on the opinions to find a suitable approach to the management of Big Data projects. Focus on the assumption that the participants in the meeting were interested in the problems of Big Data. We will not focus on a specific management approach, as we also assumed that not all participants have in the companies established agile approach to project management, and therefore may not have knowledge about it.

4.1. Use of agile approach in the management of Big Data projects

The aim of the question which is important factors in the management of Big Data projects was to determine whether the management of Big Data projects can apply the value of the agile manifesto or if at all agile approach is applicable in managing these projects (Fig. 4).

It is interesting that respondents in possibility choose two most important factors in the management of Big Data projects, naturally chose the first three partitions of values that prefer Agile Manifesto. Also considered important to the processes and tools which, according to the manifesto should have less weight than people and communication.<sup>10</sup> All values of higher priority, respondents rated more than those shown on the manifest, which are not crucial prerequisites for a successful project. In several lectures of the workshop was said recommendation to iterative approach in implementing Big Data projects. Based on the above, we can assume that the agile approach can be used to manage Big Data projects.

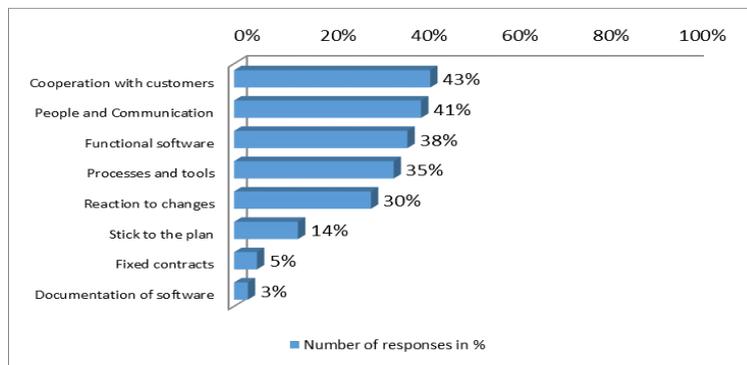


Fig. 4. What do you think is the most important in the management of Big Data projects?

4.2. Practical recommendations for the implementation of Big Data projects

General recommendations for the implementation of Big Data projects are currently not comprehensively summarized.<sup>23</sup> We wondered with which recommendations put forward by Internet respondents identify themselves as the object of this work is not only to determine what approach to take, check whether the agile approach is appropriate, but also to define the best approaches in the management of Big Data projects. Respondents were asked to select the two options to answer the question (Fig. 5): What are the recommendations in implementing Big Data solutions in organization? A frequent problem is the lack of the right skills in the team,<sup>25</sup> the second most important recommendation of the respondents considered assemble the right team (32%). As in several lectures stated, at Big Data projects is needed to involve a person such as the manager, analyst or data scientist, who will have the domain knowledge and will be able to ask the right questions. Tranık, consider as a priority responsibility role of manager to improve key performance indicators and setting realistic goals and expectations of Big Data project with which it shares 52% of responses. The analyst is responsible for the specification and quality of data to be collected. Its task is to cooperate with the IT department that collects specified information. People from IT areas usually do not have sufficient skills to ask the right questions.<sup>25</sup> Respondents also recommended removing communication barriers

between managers and development teams (30%). Between managers and developers should stand analyst. <sup>20</sup> If we're lucky, we will be able to put together technical and sales team, what is challenging, <sup>21</sup> but it ensures cooperation and can eliminate communication barriers.

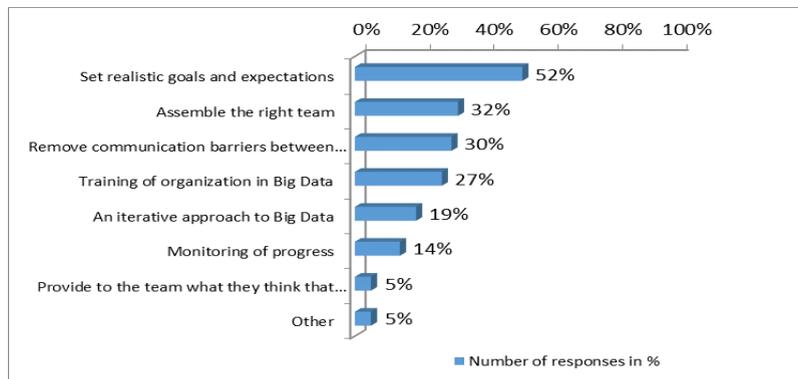


Fig. 5. What are the recommendations for the implementation of Big Data in the organization?

Buday in his lecture outlined the learning in the area of Big Data as a challenge, because on the market is still small offer of education in this area and the development of skills in most cases is left to the individuals. <sup>23</sup> 27% of responses considered the training of organization in the area of Big Data for important recommendation for the implementation of Big Data projects.

It is recommended to integrate Big Data with innovations, <sup>25</sup> so starts in small, often fail what ultimately move you forward. In the questionnaire, 19% of responses agreed that iterative approach to Big Data is one of the two most practical recommendations. Big Data is recommended to implement in an organization's environment, whereas in the company are working people who understand the area of the domain of functioning of this company. As well Kovacic said, <sup>21</sup> in terms of access to the management of Big Data projects, it is appropriate to iterate in the following steps: defining the problem, identifying the knowledge or information gap, setting hypothesis, test hypotheses (accept or reject), continue in the cycle (iteration). Defining the problem and finding the information gap best manage the people who know the domain, so workers in the organization. At the initial stage of realization it is appropriate to allocate a small team, to whom we give as task experimenting with company information. It is appropriate to use a hosted platform in order to fully concentrate on finding the questions and subsequent recommended steps of the technical implementation, <sup>25</sup> which ultimately in his presentation said the Tranik.

Further recommendations for the implementation of Big Data projects on the basis of the workshop are the following:

- Determining the scope of the project based on the number and quality of data sources that are available <sup>19</sup>
- Gaining knowledge about current data <sup>22</sup> and the identification of data sources in the future <sup>20</sup>
- Establishing safety limits of the project, based on the findings of events that should be earmarked in the analysis. <sup>22</sup>

## 5. Results

At Big Data we have large amounts of unstructured data, which we want to use to our advantage. That combines with many aspects of modern technology, but the principle is that we, as managers should be interested in return on investment. And what is related to return on investment? It may appear only if you know the right questions to ask and define the right business problems. Then we can come to the determination of hypotheses which we can accept or decline on the base of the data. <sup>21</sup> And there is great need for technological support. It is not possible that the company will define some business issues, will create Big Data project and begin analyzing possible answers. Because until the implementation of the project is completed, the questions may no longer be current at all.

If the suitability of agile and plan driven approach is determined according to the fundamental distinguishing methodologies of these approaches (Table 1), we can see that the choice of approach in the management of Big Data projects is conditioned not only by the size and criticality of the project, but also by the dynamics of the environment, the skills of people in charge of project work and organizational culture.

In case of bigger projects, which include the implementation of the overall Big Data technology and infrastructure and are critical to safety, is a more suitable plan managed approach. Agility in the management, of course, is applicable, but must be eliminated by goal of the project. Agile approach to the management of Big Data projects can be applied to smaller, less critical projects whose implementation takes place directly in the organization. The aim of these projects should be the search for possibilities of current business data to gain value from them.

Table 1. The suitability of agile and plan driven approach.

	Agile approach	Plan managed approach
<b>Size</b>	A smaller Big Data project that for the implementation uses SaaS, cloud and hosting platforms.	Bigger Big Data project with a complex Big Data technology solution.
<b>Seriousness</b>	Small projects that use SaaS from external suppliers at the beginning do not require a large financial investment in Big Data technologies and are less critical in case of failure.	Large critical project for which is an increased need for data security. These projects are expensive to implement Big Data technologies and therefore critical in terms of investment.
<b>Dynamics</b>	They are suitable for Big Data projects, which seek to use business data in a dynamic environment (telecommunications), or R & D.  Agility enables to respond to changing requirements, open questions, rephrasing of questions and determine the hypotheses, after their acceptance (resp. rejection) is continuing in the next iteration.	They are suitable for a more stable environment (eg. government) and the environment in which security is a critical factor, because it is necessary to prepare the plan.  In solving infrastructure across the enterprise, it is necessary to plan the different stages of implementation, technological solution, it is also necessary to manage the reconciliation of the technological objectives and technological solution.
<b>Staff</b>	Smaller highly specialized team on the business domain of company that focuses on asking the right questions.	Larger team of technologically specialized experts, analysts and manager, who will manage this team and monitor its progress.
<b>Culture</b>	They are suitable for Big Data projects, which the organizational structure is with a higher degree of freedom, what is allowing experimentation with the data and searching of value in the data.	They are suitable for larger companies with autocratic manner of leadership where on the technology development are involved more people.

For such a project should not be the intention to implement extensive and complex Big Data technology infrastructure that can be replaced with hosted platforms, cloud and SaaS. Using the principles of the Agile Manifesto for the management of such Big Data projects is handled in Table 2.

The close cooperation of managers, CIOs, the owners of the product, the development team can after the determination of hypotheses help find the right data, cleanse them, and they can help in the decision to adopt or reject a hypothesis. In these cases, the agile iterative approach is very important because with Big Data is difficult to pre-determined return on investment. Therefore, we recommend continuing in small steps and implement technical cooperation and people who are familiar with business issues. Although the provision of this cooperation is demanding not only on time, but also on management, it's probably the only way we can Big Data use completely in real time, with less risk, low-investment and greater guarantee of return on investment in the short term.

Table 2. Use of agile principles in the management of small Big Data projects in the organization.

1. Satisfying customers with before and continual supply of valuable software is applicable to a wider ROI.	2. Changes in requirements are welcomed, which allows to ask new emerging issues in a dynamic environment and deliver competitive advantage.	3. Frequent delivery of functional software in the shortest intervals can be applied to the laying of specific questions of business and determine hypotheses.
--	---	---

<p>4. Cooperation between analysts and developers during the project will allow analysts to identify sources of data that will be implemented during the development and consistently answer questions about the developers requirements.</p>	<p>5. Assemble motivated team and provide it with the support and confidence will allow experimentation with the data in a less bureaucratic environment.</p>	<p>6. Implicit knowledge in the team are sufficient, since it is a smaller scale with a limited focus on business issues of the organization in which they work.</p>
<p>7. Whereas it is appropriate to use SaaS, a key indicator of progress for the IT department can be managing of a hosted software, adding new data sources that are required to answer questions, or to implement corporate solution through API.</p>	<p>8. By assuming small iterations and a good knowledge of the domain of team members it can sustain a constant pace of work on the project.</p>	<p>9. When using Saas is not a key factor for the success technological knowledge. When choosing technology solutions is appropriate to assess in advance what is the support provided to them.</p>
<p>10. In these cases is appropriate to look for the simplest and most effective ways to accept or reject the hypothesis.</p>	<p>11. Smaller companies have a more decentralized style of management than corporate environment. It is possible to leave the team to organize itself to be able to create correct requirements of the project and to seek new challenges and information gaps.</p>	<p>12. It is appropriate that the team is evaluated at regular intervals to check for a return on investment of individual deliveries. This will help increase the efficiency of the team, control returns, and propose effective changes. The project may be terminated prematurely in the event of poor returns.</p>

Based on Practical recommendations for the implementation of Big Data projects we have identified the following practical recommendations:

- set realistic goals and expectations,
- determine the scope of the project based on the number and quality of data sources available to us,
- select an iterative approach to Big Data,
- assemble the right team,
- remove communication barriers between managers and those who are entrusted with the implementation,
- training of organization in the Big Data,
- knowledge of current data and identifying data sources for the future,
- establishing safety limits of the project.

At Big Data management is recommended to apply the following recommendations:

- restrict poor quality and inappropriate data,
- ensure the preparation of the information,
- take Big Data as part of corporate strategy,
- include the complexity of the data (ie. metadata),
- do not to overlook granularity of data,
- get support from management,
- define the total operating costs,
- avoid data lakes and ensure cooperation between departments,
- define return on investment,
- correct set of data into context.

## Conclusion

What approach should accordingly be applied in the management of Big Data projects? If we consider the comparison based on the Triple Constraint of the project we can determine that for management of Big Data projects is the preferable agile approach. This is because of the constantly changing requirements posed by the giving of new questions. In implementing of Big Data project, it is recommended to start with small use case, accept small failures that move us forward and continue iterative approach. The problem arises if the company does not want or cannot use hosted or cloud platform provider. The reason as safety requirements in the domain such as banking or government must be specified in detail and analyzed at the start of the project.

## References

1. Duncan WR. *A guide to the project management body of knowledge*. Project Management Institute 1996.
2. Boehm BW. *A spiral model of software development and enhancement*. *Computer*, 1988, 21.5: 61-72.
3. Satpathy T. *A guide to the Scrum body of knowledge*. 2013.
4. STN ISO 10006: 2004 *Systémy manažérstva kvality. Návod na manažérstvo kvality v projektoch*.
5. Google. *Google Trends* [online]. 2015 [cit. 2015-01-23]. Online: <http://www.google.com/trends/>.
6. Vokurka RJ.; Flidner G. *The journey toward agility*. *Industrial Management & Data Systems*, 1998, 98.4: 165-171.

7. Domino MA; Collins RW; Hevner AR. *Controlled experimentation on adaptations of pair programming*. Information Technology and Management, 2007, 8.4: 297-312.
8. One Version. *The 8th Annual "State of Agile" Survey [online]*. 2013 [cit. 2015-01-23]. Online: <http://www.versionone.com/pdf/2013-state-of-agile-survey.pdf>.
9. Hunt J. *Agile software construction*. 2006.
10. Ceschi M, et al. *Project management in plan-based and agile companies*. Software, IEEE, 2005, 22.3: 21-27.
11. Rouse M. *Big data management*. In SearchDataMnagement [online]. 2013 [cit. 2015-01-25]. Online: <http://searchdatamanagement.techtarget.com/definition/big-data-management>.
12. Accenture. *Big Success With Big Data*. Accenture Analytics survey [online]. 2013 [cit. 2015-01-25]. Online: <http://www.accenture.com/SiteCollectionDocuments/PDF/Accenture-Big-Data-POV.PDF>.
13. Stamford C. *Gartner Survey Reveals That 64 Percent of Organizations Have Invested or Plan to Invest in Big Data in 2013*. In Gartner Newsroom [online]. 2013 [cit. 2015-01-25]. Online: <http://www.gartner.com/newsroom/id/2593815>.
14. Stamford C. *Gartner Survey Reveals That 73 Percent of Organizations Have Invested or Plan to Invest in Big Data in the Next Two Years*. In Gartner Newsroom [online]. 2014 [cit. 2015-01-25]. Online: <http://www.gartner.com/newsroom/id/2848718>.
15. Sagiroglu S; Sinanc D. *Big data: A review*. In: Collaboration Technologies and Systems (CTS), 2013 International Conference on. IEEE, 2013. p. 42-47.
16. Kataria M, Mittal Ms P. *BIG DATA: A Review*. 2014.
17. Beyer MA; Laney D. *The importance of 'big data': a definition*. Stamford, CT: Gartner, 2012.
18. Laney D. *3D data management: Controlling data volume, velocity and variety*. META Group Research Note, 2001, 6.
19. Hrapko J. *My experience on Hadoop implementation and what would I do different next time*. Faculty of Management UK, Odbojárov 10, Bratislava. In Big Data Management. 2014. Interview.
20. Trančík I. *Applying agile methods to big data projects that deliver business value*. Faculty of Management UK, Odbojárov 10, Bratislava. In Big Data Management. 2014. Interview.
21. Kovačič I. *Is project management of a big data project really different?* Faculty of Management UK, Odbojárov 10, Bratislava. In Big Data Management. 2014. Interview.
22. Majerník M. *My corporate experience in management of R&D projects*. Faculty of Management UK, Odbojárov 10, Bratislava. In Big Data Management. 2014. Interview.
23. Buday T. *Big Data Management in real live or Best practices with BDM*. Faculty of Management UK, Odbojárov 10, Bratislava. In Big Data Management. 2014. Interview.
24. Kelly J, Kaskade J. *CIOs & BIG DATA What Your IT Team Wants You to Know*. In Infochimps.com [online]. 2014 [cit. 2015-04-17]. Online: <http://www.infochimps.com/resources/report-cios-big-data-what-your-it-team-wants-you-to-know-6/>.
25. Asay M. *8 Reasons Big Data Projects Fail*. In InformationWeek.com [online]. 2014 [cit. 2015-04-17]. Online: <http://www.informationweek.com/big-data/big-data-analytics/8-reasons-big-data-projects-fail/a/d-id/1297842>.
26. M. Gregus, N. Kryvinska, "Service Orientation of Enterprises - Aspects, Dimensions, Technologies", 2015, Comenius University in Bratislava, ISBN: 9788022339780.
27. N. Kryvinska, M. Gregus, "SOA and its Business Value in Requirements, Features, Practices and Methodologies", 2014, Comenius University in Bratislava, ISBN: 9788022337649.
28. M. Stoshikj, N. Kryvinska, C. Strauss, "Efficient Managing of Complex Programs with Project Management Services", Springer, Global Journal of Flexible Systems Management, Special Issue on Flexible Complexity Management and Engineering by Innovative Services, (March 2014) 15(1):25–38, DOI 10.1007/s40171-013-0051-8.
29. N. Kryvinska, "Building Consistent Formal Specification for the Service Enterprise Agility Foundation", The Society of Service Science, Journal of Service Science Research, Springer, Vol. 4, No. 2, 2012, pp. 235-269.
30. E. Molnar, N. Kryvinska, M. Greguš, "Customer Driven Big-Data Analytics for the Companies' servitization", The Spring Servitization Conference 2014 (SSC 2014), 12-14 May 2014, Aston Business School, Aston University, UK, Eds. T. Baines, B. Clegg, D. Harrison, pp. 133-140.
31. M. Stoshikj, N. Kryvinska, C. Strauss, "Project Management as a Service", The 15th International Conference on Information Integration and Web-based Applications & Services (iiWAS2013), 2-4 December 2013, Vienna, Austria, ACM, pp. 220-228.