

Information security in the context of the digital economy

P.Yu. Filyak
Syktyvkar State University
Syktyvkar, Russia
paralax-1@yandex.ru

Abstract

The article deals with an approach to solve the problems of ensuring information security in the digital economy, when, in particular, such key issues as working with Big Data and information security become one of the basic initial requirements for the "Digital Economy of the Russian Federation".

Keywords: digital economy, information security, Big Data, standards, Data Mining, visualization and exploration software, graph DataBase, information-analytical system (IAS).

1 Introduction

In July 2017, the Government of the Russian Federation adopted an extremely important document - Order of the Government of the Russian Federation No. 1632-r of July 28, 2017 "On the Approval of the Digital Economy of the Russian Federation Program" (hereinafter, the Program), which identifies strategic directions and technologies: large data; **neurotechnology and artificial intelligence**; distributed registry systems; quantum technologies; new production technologies; industrial Internet; components of robotics and sensorics; wireless technology; technology of virtual and augmented realities.

Among the challenges and threats to the development of Russia's digital economy are:

- threats to the individual, business and the state, related to trends in the construction of complex hierarchical information and telecommunication systems that widely use virtualization, remote (cloud) data storage, as well as various communication technologies and terminal devices;
- building the capacity of external information and technical impact on the information infrastructure, including critical information infrastructure;
- the growth of computer crime, including international crime.

One of the main tasks of the Program, the implementation of which is scheduled for 2024, is the formation of the Russian Federation as one of the world leaders in the field of information security. To fully implement the tasks set in the Program, one of the main initial requirements is the implementation of a set of measures to ensure the information security of the Program. In this regard, the Government Commission on the use of information technologies to improve the quality of life and the conditions for doing business has approved the "Action plan for the direction"Information Security"of the program "Digital Economy of the Russian Federation"(Minutes No. 2 of December 18, 2017).

Copyright © by the paper's authors. Copying permitted for private and academic purposes.

In: M.Yu. Filimonov, S.V. Kruglikov M.S. Blizorukova (eds.): Proceedings of the International Workshop on Information Technologies and Mathematical Modeling for Efficient Development of Arctic Zone (IT&MathAZ2018), Yekaterinburg, Russia, 19-21-April-2018, published at <http://ceur-ws.org>

What problems arise in connection with the advent of the digital economy [1-8], in addition to the ones outlined above? First of all, this is the problem of a great increase in the amount of data and information that we have to deal and operate with. The human brain capable of accommodating approximately 35 exabytes of information, cannot, in principle, "absorb and process" all the information available in the world and embrace the vast, which marks the transition beyond the barrier of perception and the onset of an era of chronically increasing information overload, when a person, (decision maker) is forced to work not with primary information, but with substrates of information, which requires the use of special approaches and methods in the development and adoption of managerial decisions. The overabundance of information can be compared with a pandemic, which is most often exponential [3].

But there appear natural questions: Does the amount of knowledge grow proportionally in conditions of obvious information overload and the number of correct administrative decisions? Is not the production of information a self-contained artificial process, self-closed, living according to the laws and on the basis of the commercial interests of the information producer? Are the volumes of information justified by the objective need of society and how do they correlate with the amount of knowledge of society? How to operate with such volumes of information/data and make management decisions in the conditions of information overload?

2 Approaches

Let's remember the pyramid of the semantic hierarchy DIKW (data, information, knowledge, wisdom - data, information, knowledge, wisdom) of Russell Ackoff proposed by him in 1989, which suggests the following gradation (Fig. 1)¹



Рис. 1: Correlation of concepts "data "information "knowledge "wisdom".

- the data layer is at the bottom;
- information adds context;
- knowledge adds "how"(the mechanism of use);
- wisdom adds "when"(conditions of use).

Thus, according to Ackoff, knowledge acts as the value of information. This is what turns information into instructions (recipes), and then these "recipes" can already be used according to one's own understanding. Therefore, we need a fine line that turns data and information into knowledge, which requires semantic dependencies (semantic metrics) and formalization, implemented in any of the adopted models - frame models, neural networks, etc. This can be visually illustrated by turning the "pyramid of Ackoff" to 180° - the pyramid will turn into a funnel. And what should we do with large amounts of data "Big Data"[7, 8]? - Apply the "funnel" and "filter" them with the help of "special filters". It acts just as the separator where milk and other dairy products are obtained from milk, discarding buttermilk (serum), or as the oil cracking (Fig. 2) producing oil products.

¹Copying κ by the paper's authors. Copying permitted for private and academic purposes.

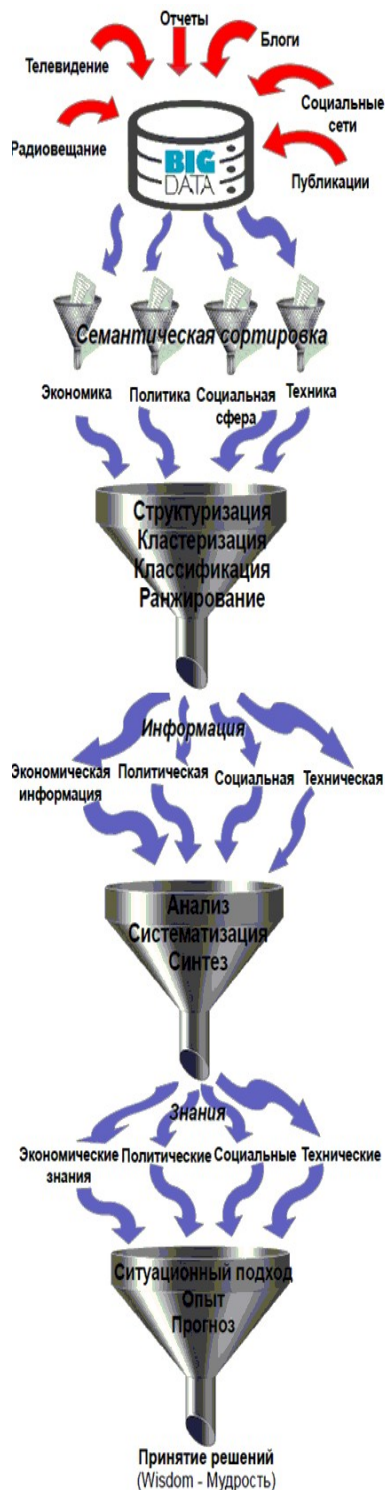


Figure 2. The Scheme of Data Mining

What and who can act as filters? Historically, the role of such filters was performed by sages, experts, priests, passing through themselves all possible volumes of information, and finally providing the user with only that truncated refined information, which in their opinion was exactly what the consumer, often very high-ranking, needed. Currently, the role of "filters" is played by teachers and parents, who "pump" the Giga- or Terabytes (the "primaries") through themselves, so that they could provide students or children with the right picture of events or a "picture of the world" in a limited time of the educational process. Under the conditions of the information society, the role of filters is performed by information and analytical systems (IAS), allowing to perform Data Mining [7,8] and providing cognitive technologies implemented in various algorithms and cognitive schemes - frame models, neural networks, artificial intelligence systems. An example of one of the effective implementations of this approach can be called the development of IBM Corporation - Watson computer; some time ago it was already a serial product of this company, involving the configuration of this computer, its machine learning and training with the accumulation of statistics, as well as intellectual assistants of different levels.

3 A little bit about Big Data and Data Mining

At present, when the volumes of data and information (Big Data) that have to be handled by decision-makers have significantly exceeded the values that the human brain can process [7, 8], as mentioned above, it is necessary to put into practice approaches that allow timely and effectively extract the knowledge from huge amounts of structured and unstructured data through intelligent analysis of data and information - Data Mining [4,5,7,8], which cannot be imagined without widespread adoption and the use of modern information and analytical systems (IAS), representing an effective symbiosis of mathematics, modern means of computer technology, software and multimedia.

At the moment, a wide range of information and analytical systems (IAS), of both domestic and foreign developers, allowing to work with Big Data and Data Mining, is available on the market, for example, the environment for visualization and analysis of Gephi data and graphs (Figure 3), graph databases Neo4j (Figure 4).

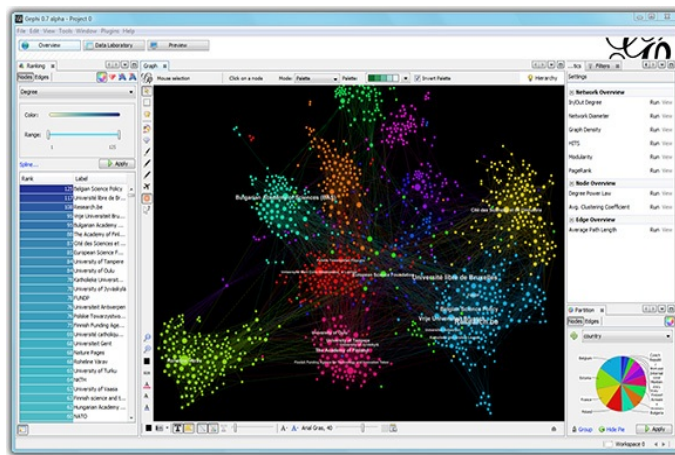


Figure 3: Data and graph visualization and analysis **Gephi**.



Figure 4: Ontologies of ensuring the legal field of information security in **Neo4j**.

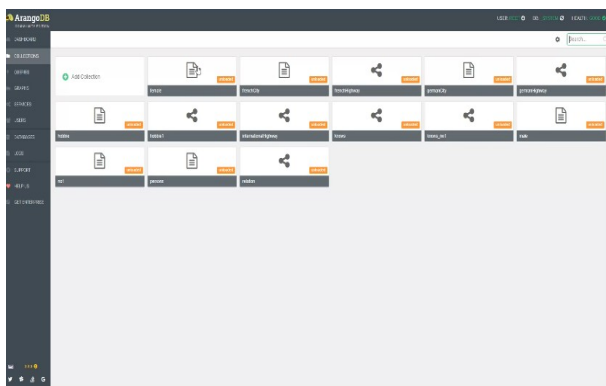


Figure 5: Matrix for establishing links in the graph database **ArangoDB**.

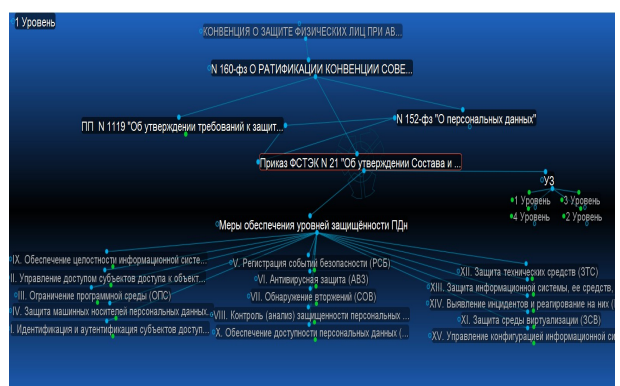


Figure 6: Ontologies of personal data security in **Personal Brain**.

ArangoDB (Fig.5), AllegroGraph and others; the knowledge management system (knowledge management technology) Personal Brain (Fig.6) and others can also be considered; information-analytical systems - Semantic archive, products aimed at implementation of Data Mining methods of the Russian developer - IAS PolyAnalyst (Fig.7), used in more than 20 countries of the world, RapidMiner (Fig.8) as an open platform solution - the environment for working with Big Data and analytics and others.

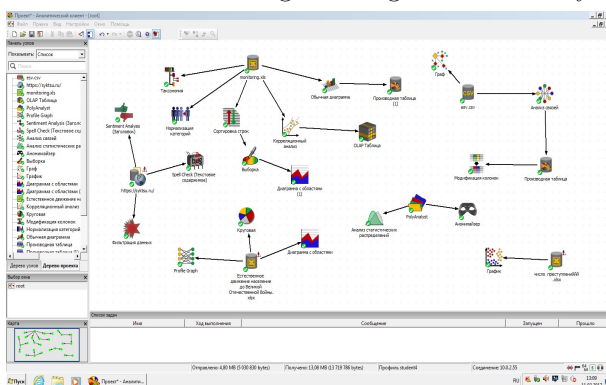


Figure 7: Data Mining in **PolyAnalyst**.

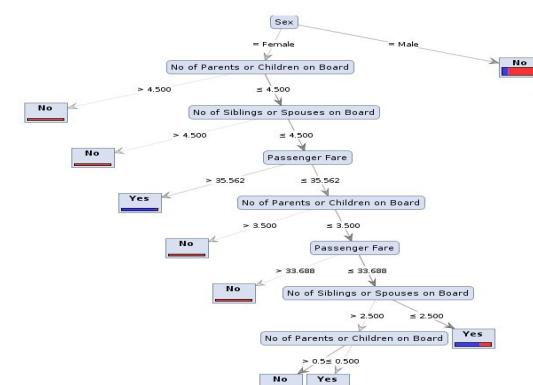


Figure 8: Decision tree built using **RapidMiner**.

The task is to select from a wide range of information and analytical systems (IAS) such a tool that would effectively achieve the assigned goals and solve specific problems at the most optimal value of the traditional criterion - the price / quality ratio, that would not have excessive requirements for the hardware/software platform, and would have the appropriate interface and ergonomics, the convenience of working with "input" and

"output"information (and its presentation).

It is extremely important to establish links and dynamic change of relationships, since information overload caused by the abundance of documents, each of which must be considered and many of which can contradict each other, is a real headache for those who are obliged to undergo certification procedures, licensing, accreditation.

First of all, this concerns the sphere of ensuring security and information security, in particular. As an example, at the moment there are more than 500 standards and normative documents on information security in the world [9]. It is very difficult to get acquainted thoroughly with these documents so that they could be meaningfully applied and, moreover, freely operated with. Only international standards in the field of information security (IB) number more than 100. In Russia, currently, there are more than 40 state standards, and about 70 regulatory acts in the field of information security and almost the same number is under development.

The author of the article presents examples of information and analytical systems (IAS), the direct use of which showed high efficiency, when working with large data (Big Data) and the implementation of Data Intelligence (Data Mining).

4 Afterword

It is impossible to solve the tasks in the digital economy without the introduction and development of such a direction as Data Intelligence. The concept itself of Data Intelligence - intelligent data processing - means:

- Collection and storage of data; it operates not only with the concept of Big Data, but also with such a fairly new term as "Data Lakes";
- Data processing and analysis;
- Data visualization and decision making on the basis of modern information technologies including cloud solutions (Cloud Computing).

It is not difficult to guess that all this powerful pyramid of a new trend with all its components is based on the "thin leg" of applied mathematical methods and imitation (mathematical) models, the entire superstructure of which depends on the adequacy, accuracy and limits of applicability, whatever capabilities it may have been announced and whatever interface, means and methods of input and means of multi-media (n-D) - presentation of the output information it may have been equipped with.

References

- [1] A.G. Ostapenko, E.A. Schvartsoph, D.A. Savinov, E.V. Gusev, D.V. Gusev. Programm Tool for modelling epidemiological processes in the social networks. *Information @ Security*. 20(1-1(4)):39-48, 2017 = А.Г. Остапенко, Е.А. Шварцкопф, Д.А. Савинов, Е.В. Гусев, Д.В. Гусев. Программный комплекс моделирования эпидемических процессов в социальных сетях. *Информация и безопасность*. 20(1-1(4)):39-48, 2017.
- [2] G. Greph. *Century of IT-specialists has finished*. https://www.gazeta.ru/tech/news/2017/10/20/n_10717526.shtml?utm_source=gazeta-top-article&utm_medium=news_readable&utm_campaign=desktop = Г. Греф. "Век айтишников" закончился. https://www.gazeta.ru/tech/news/2017/10/20/n_10717526.shtml?utm_source=gazeta-top-article&utm_medium=news_readable&utm_campaign=desktop.
- [3] V.I. Budzko, A.V. Schmid, N.A. Ivanov, S.P. Sadovsky. *Security @ confidentiality of a centralized system of disease prevention*. <https://yadi.sk/d/F-CaJt3J3DgeS4> = В.И. Будзко, А.В. Шмид, Н.А. Иванов, С.П. Садовский. *Безопасность и конфиденциальность централизованной системы профилактики заболеваний*. <https://yadi.sk/d/F-CaJt3J3DgeS4>.
- [4] V. Mayer- Shönberger, K. Kukier. *Big Data. A revolution that will change the way we live, work and think*. John Wiley & Sons, 2013.
- [5] *Big Data: changing the future of humanity*. http://www.kitaichina.com/se/txt/2013-03/28/content_530849.htm = *Большие данные (Big Data): изменение будущего человечества*. http://www.kitaichina.com/se/txt/2013-03/28/content_530849.htm
- [6] N.V. Grishina. *Organization of complex information protection system*. М.: Helios ARV, 2007 = Н. В. Гришина Н. В. *Организация комплексной системы защиты информации*. М.: Гелиос АРВ, 2007.

- [7] P.Yu. Filyak. Design under condition security information provision. *Information @ Security*. 18(1):101-106, 2015 = П.Ю. Филяк. Проектирование с учетом обеспечения безопасности. *Информация и безопасность*. 18(1):101-106, 2015.
- [8] P.Yu. Filyak. Information security and complex information security system: analysis approaches. *Information @ Security*. 19(1):72-79, 2016 = П.Ю. Филяк. Информационная безопасность и комплексная система безопасности: анализ, подходы. *Информация и безопасность*. 19(1):72-79, 2016.
- [9] A. M. Astakhov. *Art information risk management*. М.: DMK Press, 2010. = А.М.Астахов. Искусство управления информационными рисками. М.: ДМК Пресс, 2010.
- [10] *The Open Graph Viz Platform. Official Developer Portal*. <https://gephi.org/>
- [11] *About GEPHI: Altsoph's Journal*. <http://altsoph.livejournal.com/169571.html>
- [12] *NEO4J. Official site of Neo4j*. <https://neo4j.com>
- [13] *ArangoDB. Official site of ArangoDB*. <https://www.arangodb.com>
- [14] *THE BRAIN. Official site of BRAIN*. <http://www.thebrain.com>
- [15] Portal of Megaputer Company. (developer of the IAS PolyAnalyst). <http://www.megaputer.ru/polyanalyst.php>
- [16] *RapidMiner - Official Developer Portal*. <https://rapidminer.com>
- [17] *RapidMiner Studio Manual*. RapidMiner, 2014.