Application of Open Data in Accordance With Information Security Requirements*

Aleksandr V. Dorofeev[1], Alexey S. Markov[2][0000-0003-0111-7377], and
Valentin L. Tsirlov [2][0000-0003-2657-4179]

[1] NPO Echelon, Moscow, Russia
`a.dorofeev@npo-echelon.com`
[2] Bauman Moscow State Technical University, Moscow, Russia
`{a.markov, v.tsirlov}@bmstu.ru`

Abstract. The paper discusses the existing inconsistency between information security requirements and requirements for government data accessibility. The paper provides the classification of open and public government data. It specifies generally accessible data sources that can be used for information security management. Completes and adequate measures to protect open resources in public information systems in Russia are provided. A conclusion is made on the adequacy and completeness of taxonomies in the area of web resource security. The authors conducted a comparative analysis of individual Internet portals on open data. The paper points out information security problems in relation to open data usage and some ways to solve them.

Keywords: Open Government Data, Public Information Systems, Publicly Available Information.

1 Introduction

The development of the paradigms of electronic and open governments has brought attention to so-called open government data (hereinafter - OD) [6, 10, 11, 15, 18], which should be made available to the public by the Internet portals of various systems of government authorities for further processing in various information systems. Due to this, two factors of society informatization should be distinguished:

- One of the ways to improve public administration is the introduction of some principles of its openness, including the placement and maintenance of OD, some of which provides a feedback loop for public control;
- Open technological movements continue to be developed, ranging from the use of open-source software products to open concepts of compliance assessment (e.g. Common Criteria), the goal of which is to make the research and production of IT products and systems more effective.

- As for OD processing, the following can be mentioned:
- Regulatory requirements for the integrity and accessibility of shared public information resources are being developed in the country;
- Confronting threats of denial of service (accessibility of resources) in cyberspace is the most difficult task to ensure the safe operation of Internet portals.

The above determines the scientific interest in the notion of OD information security itself. It should be noted than in terms of everyday life, from the point of an outdated view of ensuring information security (as securing privacy and confidentiality of the information), an open data security procedure seems to come into apparent conflict with the very notions of openness and general accessibility. However, this is certainly not the case, if we consider the notion of information security in the generally accepted international understanding, when all open resources of information systems can have a wide range of threats in the area of information, primarily with respect to threats to integrity and accessibility.

The research of the OD information security assurance is the subject matter of this paper.

2 The notion of open data and classification

Open government data normally include publicly available data in electronic format, officially provided by government authorities for further free use. For example, the server of government authorities of the Russian Federation provides these data as information on the activities of government authorities and local governments placed on the Internet in a format that allows its automatic processing for reuse without any prior human modification (machine-readable format) and can be freely used in any lawful purposes by any persons irrespective of the form of its placement.

The basic principles of OD are identified, such as primacy, completeness, relevance, machine readability, lack of discrimination on access, lack of proprietary formats, clean license, etc. [1, 22, 24].

In technical terms, OD sets have the property of interoperability, since such data should be free of any access or implementation restrictions, and have open-ended formats and interfaces [8].

According to the guidelines of the Russian Federal Agency for Scientific Organizations, the types of open data are classified by the following main criteria: data domain, data format, a data structure (linear, hierarchical, etc.), data volume, publication method, storage method, updating frequency, data relevance.

As mentioned above, the notion of OD is defined by attributing data sets to the governmental resource subject to placing them on the Internet for general use in the specified formats (see Fig 1). Thus, OD is a subset of public government data (PD) characterized by additional restrictions, mainly in interfaces and presentation format on the Internet portal:

$$OD \underset{def}{\subseteq} PD. \tag{1}$$

At the same time, there are no technical obstacles both for converting public data to the open data format and for further conversion of OD.

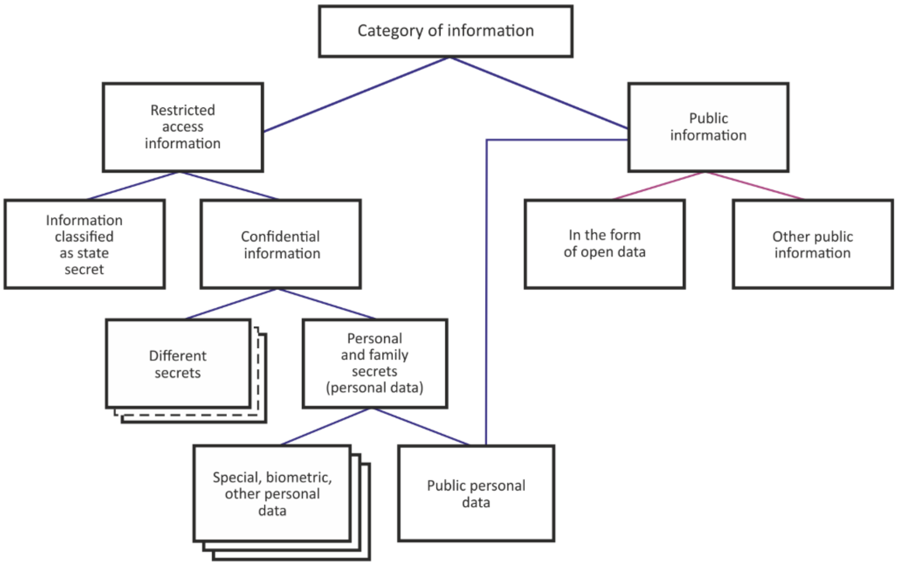

Fig. 1. Classification of information for accessibility.

It is easy to see (Fig. 1) that in terms of legislation there are two categories of public data covered by the state information security (IS) requirements:

- Public personal data in personal data processing information systems (PDIS);
- Public data in public information systems (PIS), which are actually OD (see below).

In terms of purposes, two subclasses of public data sets in the area of information security can be distinguished:

1. OD sets related to the official activities of public authorities;
2. Other sets of public government data which can be used in the IS management systems.

As for public authorities, there are three regulators in the area of IS in the Russian Federation. The most useful and complete IS information can be obtained on the official Internet portal - www.fstec.ru, in the Open Data section (Table 1).

Table 1. Open data sets of information security.

No	Data set description	Formats
1	State register of certified information security means	ods, csv
2	List of testing laboratories	ods, csv
3	List of appraisal bodies	ods, csv
4	List of certification bodies	ods, csv
5	List of subordinate organizations	csv
6	List of local agencies	csv
7	Plan of scheduled inspections on licensing control issues	ods, csv
8	Plan of scheduled inspections on export control issues	ods, csv
9	Register of information security means licenses	ods, csv
10	Register of technical protection of confidential information licenses	ods, csv

You can certainly find other publicly available IS data on the IS portal, such as the register of expert organizations, register of educational institutions and centers, regulations, guidelines, specific, anti-corruption, tender information, etc. Particular attention should be certainly paid to the Information Security Threats Data Bank.

Other sets of publicly available data can easily be obtained using competitive intelligence methods [4] from publicly available government sources (Table 2).

Table 2. Examples of publicly available government sources.

Data sets, documents	Examples of Sources
Guidance support of activities	
IS standards, guidelines	www.gost.ru/
Theses, peer-reviewed publications and etc.	http://vak.ed.gov.ru/dis-list
Services, supplies	
Unified register of Russian programs	https://reestr.minsvyaz.ru/
Register of certifying centers	http://minsvyaz.ru/ru/activity/...
Checking joint contractors of work	
Register of unscrupulous suppliers	http://fas.gov.ru/
The state registration of legal entities and etc.	https://egrul.nalog.ru/
Personnel check	
Personal data of public officers	Portals of state bodies
Litigations and etc.	Court portals, e.g.: http://mos-gorsud.ru
Planning of investments to the regions	
A budget of the constituent entity of the RF	http://budget.gov.ru
Various generalized statistics	www.gks.ru

It should be noted that the use of OD and open data of non-governmental corporations and associations give a powerful synergetic effect [5].

3 Threats to open data in the information environment

Information security of open government data can be defined as a property of OD security against threats in the information area. These threats can be divided into possible violations of technical requirements to OD and data abuse [20, 27]. In the legal framework, three current areas of OD threats are normally pointed out:

1. IS threats related to the OD life cycle and organizational and technical support systems for OD;
2. Threats related to the protection of personal and family secrets or privacy (personal data);
3. Threats related to national (state) security.

Threats related to national security are certainly the most controversial. As is known, open resources are the main source of modern intelligence, regardless of the fact whether it refers to business intelligence or has a national status. The facts of using the potential of open data for criminal purposes are fairly well known [12, 17]. OD, among other things, can be used to robotize the process of collecting and conducting a cognitive analysis of intelligence data. These issues are described in the professional literature and even in the standards [14].

Threats related to private secrets refer to the protection of the rights of subjects of personal data, which is now sufficiently developed [9, 13, 19, 23]. The intersection of interests can occur when disclosing, say, the incomes of public officers for anti-corruption purposes.

As for IS, we can single out the following threats:

- Threats to the integrity and accessibility of OD themselves;
- Threats to information security (integrity, accessibility, and confidentiality), which are associated with support systems (information, software, technical support, etc.) for PIS.

For example, if we imagine that documents on the website are recorded in the OD format, then there are threats to their integrity and accessibility. If a web portal has the functions of registering users and supporting correspondence, this requires the additional protection of this information against the disclosure threats. Similar tasks need to be solved for the protection of internal portal structure when the access to the resources is differentiated once the administrator and user privileges are differentiated. Modern PIS are connected to the interagency electronic interaction system (IEIS) and support the Unified Identification and Authentication System (UIAS), and some of them also include payment components. It should be realized that modern OD can have a distributed form (for example, link open data [1, 8, 21, 28]), - this requires that not only physical but also logical (semantic) integrity of open resources, etc. should be maintained.

Next, we consider regulatory requirements for these systems.

4 Regulatory information security requirements

It is fair to say that Russia has a regulatory framework both for informatization objects (information systems) that process OD, and for the OD protection means [2].

Requirements for information security in information systems that process OD are established by the Decree of the Government of the Russian Federation dated May 18, 2009 N 424. According to this order, two classes of public information systems (PIS) are introduced.

Currently, for the protection of PIS, several classes of information security tools have been defined, namely: cryptographic protection means, antivirus, firewalls, intrusion detection systems, and access control systems.

As to the requirements for the information security tools (except for cryptographic), the FSTEC of Russia is currently preparing them based on the same open methodology Common Criteria [2]. Nowadays, PIS-2 uses information security means of the *4th security class*, which corresponds to the 3rd increased Evaluation Assurance Level - *EAL3+* (Table 3).

Table 3. List of special regulations on information security tools.

Information security tools	Guidelines, Security Profiles
Antivirus tools	IT.SAVZ.A4.SP, IT.SAVZ.B4.SP, IT.SAVZ.V4.SP, IT.SAVZ.G4.SP
Firewalls	IT.ME.A4.SP, IT.ME.B4.SP, IT.ME.V4.SP, IT.ME.G4.SP, IT.ME.D4.SP
Intrusion detection systems	IT.SOV.S4.SP, IT.SOV.U4.SP
Access control tools (ACT)	RD SVT (1992)
Operating system with built-in ACT	IT.OS.A4.SP, IT.OS.B4.SP, IT.OS.V4.SP

It should be understood that the FSTEC of Russia additionally determines requirements (in terms of non-cryptographic methods) for state information systems (SIS) and for personal data processing information systems (PDIS) (see Table 4.). These orders adhere to a quasi-risk-based approach, i.e. they allow a reasonable reduction in the number of security measures, depending on the limited IT architecture of the system.

Currently, verification of vulnerabilities is a mandatory procedure for the design and certification of secure systems, as well as the certification of information security tools. Therefore, to give a complete picture, we briefly treat a security control issue for web portals as a basic platform to organize PIS.

Table 4. Organizational and technical measures to protect the information in information systems.

Group of organizational and technical measures	SIS	PDIS
Identification and authentication of access subjects and objects	+	+
Access control of access subjects to access objects	+	+
Restriction of the software environment	+	+
Protection of machine-readable media	+	+
Security event logging	+	+
Antivirus protection	+	+
Intrusion detection	+	+
Control (analysis) of information security	+	+
Ensuring the integrity of the information system and information	+	+
Ensuring the accessibility of information	+	+
Protection of the virtualization environment	+	+
Protection of technical means	+	+
Protection of the information system, its means, communication, and data transmission systems	+	+
Incident detection and response	-	+
Management of the configuration of the automated control system and its protection system	-	+

5 Vulnerabilities and attacks on web portal resources

Web portals are currently most often exposed to computer attacks from various types of violators. So, according to WhiteHat Security, the percentage of computer attacks on web resources is 40 % of all registered attacks over the past year [26].

An example of a typical secure web portal is shown in Fig. 2. The figure shows the main components of the web portal (web server, user application server, database management system, including OD), and standard information security means. Despite the security tools (for a known reason of extreme complexity and dynamism of software subsystems), the web portal needs constant checking for vulnerabilities and checking the threats of their implementation (the possibility of computer attacks) [29].

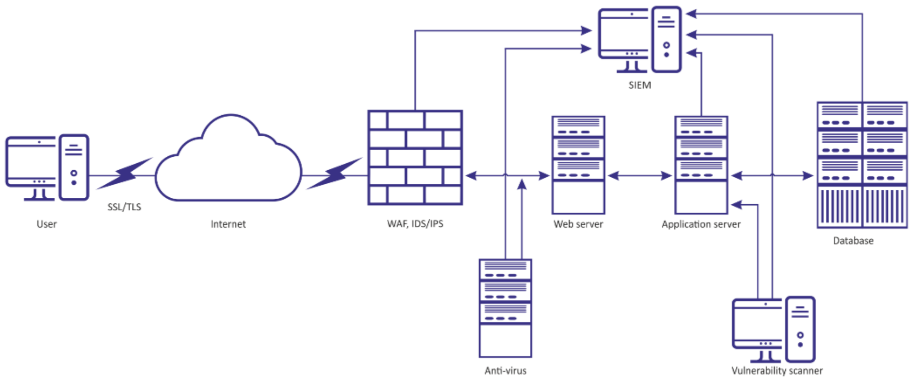

Fig. 2. Example of a secure web portal scheme.

However, similarly to the Russian regulatory framework (see above), it can be stated that the expertise of the web portal security has been globally developed, and there are relevant registers and standards for describing vulnerabilities, threats and attacks on web resources.

First of all, we can mention the classification of the Web Application Security Consortium on security threats and attacks on web resources - WASC Threat Classification [25], statistics of the OWASP open project on the ten most dangerous violations and attacks on web applications - OWASP Top 10 Application Security Risks [16], as well as the templates of attacks of MITRE organization – CAPEC [3], including on-web resources.

The Information Security Threats Database (TDB) of the FSTEC of Russia currently supports the known threats to web resources and contains the actual vulnerabilities of relevant applications and platforms [7]. An example of the TDB threats and vulnerabilities fragment of the FSTEC of Russia is shown in Table 5.

Table 5. Examples of threats and vulnerabilities descriptions of web resources in the TDB of the FSTEC of Russia.

No.	Name of threats or vulnerabilities
UBI.041	The threat of cross-site scripting.
UBI.042	The threat of cross-site forgery of the request.
BDU:2015-10929	The vulnerability of the Apache HTTP Server, which allows an intruder to bypass existing access restrictions.
BDU:2016-00707	The vulnerability of the Nginx proxy server, which allows an intruder to cause a denial of service.
BDU:2016-00484	The vulnerability of the SAUTER module Vision controller visualization web server via BACnet/IP networks, which allows an intruder to obtain confidential information.
BDU:2016-00258	Vulnerabilities of IniNet Solutions GmbH's SCADA Web Server, which allows an intruder to implement arbitrary code.
BDU:2016-00483	The vulnerability of the SAUTER module Vision controller visualization web server via BACnet/IP networks, which allows an intruder to introduce arbitrary web or HTML code.

6 Conclusions

Based on the analysis of the open government data use in the area of information security in the Russian Federation we can draw the following conclusions:

1. OD is an example of open technological initiatives aimed not only at improving the efficiency of public administration but also at improving the efficiency of research and production of products and systems, as well as in the area of information security.
2. In technical terms, OD allows us not only to ensure compliance with the key management requirement in the area of information security (according to ISO/IEC 27001) in terms of *awareness* but also to automate this process through public interfaces and formats.
3. The main threats to OD are the threats to accessibility and integrity, but the government web portals have a wide range of modern threats to information, and the number of computer attacks on web portals is steadily growing.
4. The development tendency of a regulatory framework of modern information processing systems for OD and OD security means, as well as the standards and registers of threats and vulnerabilities of web resources, is generally determined, which facilitates the activities of developers of these systems.
5. At present, OD for IS (with the possible exception of the publicly available data of the FSTEC of Russia) is presented in Russian Cyberspace very modestly. For example, there is no OD relating to the work of state security operation centers yet, and even there is no territorial statistics on cybercrime and cybersecurity.
6. It can be assumed that the use of formatted publicly available data in IS will develop. The objective reason for this is the emergence of new information conveniences

(hence, economic benefits), accompanying the informatization of society and the formation of secure virtual cyberspace.

References

1. Attard, J., Orlandi, F., Scerri, S., Auer, S.A.: Systematic Review of Open Government Data Initiatives. Government Information Quarterly. 32(4), pp. 399-418 (2015). DOI: 10.1016/j.giq.2015.07.006.
2. Barabanov, A., Markov, A.: Modern Trends in The Regulatory Framework of the Information Security Compliance Assessment in Russia Based on Common Criteria. In: The 8th International Conference on Security of Information and Networks (SIN '15), pp. 30-33. ACM New York (2015). DOI: 10.1145/2799979.2799980.
3. Common Attack Pattern Enumeration and Classification: A Community Resource for Identifying and Understanding Attacks, The MITRE Corporation, https://capec.mitre.org, last accessed 2019/05/05.
4. Dorofeev, A., Markov, A., Tsirlov, V.: Structured Approach to the Social Network Analysis of Information about a Certain Individual. In: 2nd International Conference on Electronic Governance and Open Society: Challenges in Eurasia, pp. 174-178. ACM, New York (2015). DOI: 10.1145/2846012.2846017.
5. Dorofeev, A.V., Markov, A.S., Tsirlov, V.L.: Social Media in Identifying Threats to Ensure Safe Life in a Modern City. Communications in Computer and Information Science. 674, pp. 441-449 (2016). DOI: 10.1007/978-3-319-49700-6_44.
6. Hohlov, Y., Styrin, E.: E-government in Russia: Strategies of formation and development. In: Global Strategy and Practice of E-Governance: Examples from Around the World, pp. 286-303. IDI Publishing, Hershey (2011). DOI: 10.4018/978-1-60960-489-9.ch016.
7. Information Security Threats Database of the FSTEC of Russia, FSTEC of Russia, http://bdu.fstec.ru/threat, last accessed 2019/05/05.
8. Janssen, M., Estevez, E., Janowski, T.: Interoperability in Big, Open, and Linked Data-Organizational Maturity, Capabilities, and Data Portfolios. Computer. 47, 10, pp. 44-49 (2014). DOI: 10.1109/MC.2014.290.
9. Kagawa, T., Saiki, S., Nakamura, M.: Developing personalized security information service using open data. In: 2017 18th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD), pp. 465-470. IEEE (2017). DOI: 10.1109/SNPD.2017.8022763.
10. Kokkinakos, P., Koutras, C., Markaki, O., Koussouris, S., Trutnev, D., Glikman, Y.: Assessing Governmental Policies' Impact through Prosperity Indicators and Open Data. In: Proceeding EGOSE '14 Proceedings of the 2014 Conference on Electronic Governance and Open Society: Challenges in Eurasia, pp. 70-74. ACM, NY (2014). DOI: 10.1145/2729104.2729134.
11. Koznov, D.V., Andreeva, O., Nikula, U., Maglyas, A., Muromtsev, D.I., Radchenko, I.A.: Survey of Open Government Data in Russian Federation. In: IC3K 2016 - Proceedings of the 8th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management 8, pp. 173-180 (2016). DOI: 10.5220/0006049201730180.
12. Larsen, H.L., Blanco, J.M., Pastor, R., Yager, R.R. (Eds.): Using Open Data to Detect Organized Crime Threats. Factors Driving Future Crime. Springer (2017). DOI: 10.1007/978-3-319-52703-1.

13. Meijer, R., Conradie, P., Choenni, S.: Reconciling Contradictions of Open Data Regarding Transparency, Privacy, Security and Trust. Journal of Theoretical and Applied Electronic Commerce Research. 9(3), pp. 45-58 (2014). DOI: 10.4067/S0718-18762014000300004.
14. Mendel, T., Blanton, T.S., Wadham, J., and etc.: National Security and Open Government: Striking the Right Balance. Preface by Alasdair Roberts. Campbell Public Affairs Institute. Maxwell School of Citizenship and Public Affairs, Syracuse University (2003).
15. Olifirov, A.V., Makoveichuk, K.A., Zhytnyy, P.Y., Filimonenkova, T.N., Petrenko, S.A.: Models of Processes for Governance of Enterprise IT and Personnel Training for Digital Economy. In Proceedings of 2018 17th Russian Scientific and Practical Conference on Planning and Teaching Engineering Staff for the Industrial and Economic Complex of the Region, IEEE, PTES, pp. 216-219 (2018). DOI: 10.1109/PTES.2018.8604166.
16. OWASP Top 10-2017 Application Security Risks. The Open Web Application Security Project (2017), https://www.owasp.org/index.php/Top_10-2017_Application_Security_Risks, last accessed 2019/05/05.
17. Pastor, R., Blanco, J.M.: The ePOOLICE Project: Environmental scanning against organised crime. European Police Science and Research Bulletin. 16, pp. 1-19 (2017).
18. Petrenko, S.A., Makoveichuk, K.A., Chetyrbok, P.V., Petrenko, A.S.: About Readiness for Digital Economy. In Proceedings of the 2017 IEEE II International Conference on Control in Technical Systems, IEEE, CTS, pp. 96–99 (2017). DOI: 10.1109/CTSYS.2017.8109498.
19. Pingo, Z., Narayan, B.: When Personal Data Becomes Open Data: An Exploration of Lifelogging, User Privacy, and Implications for Privacy Literacy. Lecture Notes in Computer Science. 10075, pp. 3-9 (2016). DOI 10.1007/978-3-319-49304-6_1.
20. Priyadarshy, S.: Big data, smart data, dark data, and open data: eGovernment of the future. In: 2015 Second International Conference on eDemocracy & eGovernment (ICEDEG), p. 16. IEEE (2015). DOI: 10.1109/ICEDEG.2015.7114483.
21. Radchenko, I., Sakoyan, A.: The View on Open Data and Data Journalism: Cases, Educational Resources and Current Trends. Communications in Computer and Information Science. 436, pp. 47-54 (2014). DOI: 10.1007/978-3-319-12580-0_4.
22. Sashinskaya, M.: Open Data: All You Want to Know About Open Data (Big Data, Transparency, Urbanism, Transportation, Sustainable Cities, Innovations, Smart Governance, e-Government). CreateSpace Independent Publishing Platform (2017).
23. Scassa, T.: Privacy and Open Government. Future Internet. 6, 2, pp. 397-413 (2014). DOI: 10.3390/fi6020397.
24. Tauberer, J.: Open Government Data: 2nd Ed. Kindle E-Book (2014).
25. WASC Threat Classification. V. 2.00 (1.01.2010), Web Application Security Consortium (2010). http://projects.webappsec.org/f/WASC-TC-v2_0.pdf, last accessed 2019/05/05.
26. Web Applications Security Statistics Report, WhiteHat Security (2016), https://www.whitehatsec.com/info/website-stats-report-2016-wp/, last accessed 2019/05/05.
27. Xu, L., Jiang, C., Wang, J., Yuan, J., Ren, Y.: Information Security in Big Data: Privacy and Data Mining. IEEE Access. 2, pp. 1149-1176. IEEE (2014). DOI: 10.1109/ACCESS.2014.2362522.
28. Zhukov, V., Komarov, M.: Semantic Control Method of the Internet of Things Based on Linked Open Data. In: 2017 IEEE 19th Conference on Business Informatics (CBI). 02, pp. 1-4. IEEE (2017). DOI: 10.1109/CBI.2017.5.
29. Zubarev, I., Radin P.: The Basic Information Security Threats in the Virtual Environments and Cloud Platforms. Voprosy kiberbezopasnosti *[Cybersecurity issues]*. 2 (3), pp. 40-45. (2014). [In Russ].