

PERSONAL DATA USAGE IN THE CONTEXT OF THE DEVELOPMENT OF SEMANTIC TECHNOLOGIES: LEGAL ISSUES

Darius Amilevičius*

*Mykolas Romeris University, Faculty of Public Security,
Department of Humanities
Putvinskio g. 70, LT-44211 Kaunas
Phone: (370 37) 303 664
E-mail: d.amilevicius@mruni.eu*

Birutė Pranevičienė**

*Mykolas Romeris University, Faculty of Public Security,
Department of Law
Putvinskio str. 70, LT-44221 Kaunas
Phone: (370 37) 303 655;
E-mail: praneviciene@mruni.eu*

Annotation. The article presents the consideration about the right to privacy of the individual vs. public interest in the context of the challenges of globalization of personal data ecosystem, new trends of information technologies and the digital economy boost. A new computing and information sharing architecture and technologies has emerged during the past 10 years. The proposed European Commission Data Protection Regulation currently under discussion by the European Council and Parliament is the most comprehensive attempt to establish new norms for the flow of personal data. While different views were expressed throughout the dialogue on these proposals and the underlying principles, it was clear that these rules when agreed would be the big step forward creating and implementing the new Personal Data ecosystem. While the complexity of operating in a decentralized and distributed networked environment poses new challenges, ensuring Personal Data security remains crucial.

Keywords: Personal Data Protection, Semantic Technologies, Right to Privacy

INTRODUCTION

21st century is sometimes referred to as “the Age of Cyberspace”¹. During the past decades human life has changed extremely. World became complex, hyper connected, and increasingly driven by insights derived from *big data*², and the rate of change shows no sign of slowing, nor does the volume of data show any sign of decreasing. But, the economic and social value of big data does not come just from its quantity. It also comes from its quality – the ways in which individual bits of data can be interconnected to reveal new insights with the potential to transform business and society. Technology and data by themselves are neutral. It is their use that can both generate great value and create significant harm, sometimes

¹ Biegel, S. *Beyond Our Control? Confronting the Limits of Our Legal System in the Age of Cyberspace*, The MIT Press, Cambridge, Massachusetts, 2001

² *Big data* is a collection of data sets so large and complex that they become difficult to process using available database management tools or traditional data-processing applications.

simultaneously. This requires rethinking traditional approaches to data governance, particularly to shift from focusing away from trying to control the data itself to focusing on the uses of data. It is up to the individuals and institutions of various societies to govern and decide how to unlock the value – both economic and social – and ensure suitable protections. In modern world a new approach to personal data is needed that is flexible and adaptive to encourage innovation, but also protects the rights of individuals. Notice and consent need to be reconsidered to be equipped for this changing world. The dialogue on rethinking personal data has a multiple stakeholder nature and must involve representatives of various legal, social, commercial, governmental and technical sectors, who shared their views on the changes occurring within the personal data ecosystem and how these changes affect the collective ability to uphold core principles. The dialogue also addressed key regional legislative and policy approaches, particularly in the context of the reform of Data Protection Regulation, proposed by the European Commission³. The global dialogue centered on a set of foundational principles that are familiar across a broad range of cultures and jurisdictions⁴.

Issues of accountability, protection, security and the overall control of personal data remain central to the global ecosystem of personal data. While the complexity of operating in a decentralized and distributed network environment poses new challenges, ensuring personal data security remains crucial.

The problem of the research: right to privacy of the individual vs. public interest in the context of the challenges of globalization of personal data ecosystem, new trends of information technologies and the digital economy boost.

The objective of the research: To analyze national Lithuanian legal framework of personal data protection and to disclose its dysfunctions in the use of personal data in the new personal data ecosystem and in the semantic technologies.

The article aims to:

- present the analysis of the legal regulation of the personal data protection in Lithuania;
- analyze trends and evolution of the modern personal data ecosystem;
- present the role, function and importance of personal data in semantic technologies.

³ In 2012, the Commission proposed a major reform of the EU legal framework on the protection of personal data. The new proposals will strengthen individual rights, tackle the challenges of globalization and new technologies, comprehensive reform of the EU's 1995 data protection rules to strengthen online privacy rights and boost Europe's digital economy, because technological progress and globalization have profoundly changed the way our data is collected, accessed and used. [http://ec.europa.eu/justice/newsroom/data-protection/news/120125_en.htm]

⁴ World Economic Forum, Unlocking the Value of Personal Data: From Collection to Usage, [interactive] [accessed 2013-05-02]

<http://www3.weforum.org/docs/WEF_IT_UnlockingValuePersonalData_CollectionUsage_Report_2013.pdf>

Methodology of the Research: Methods of systemic and analytical-critical analysis were applied for the research of the article. Methods of documentary analysis, meta-analysis and generalization were used as well.

LEGAL REGULATION OF PERSONAL DATA USAGE AND PROTECTION IN LITHUANIA

First instance of the Law on Legal Protection of Personal Data Protection of the Republic of Lithuania (hereinafter LLPPD) was adopted in June 11 1996⁵. It had only 14 articles establishing some general principles for data handling responsibilities and regulations regarding the rights of data subjects

LLPPD substantially was revised in 2000, and the new version of the law was adopted in July 17 2000⁶. It was expanded to 29 articles, and its purpose was to implement the Framework of the European Communities of the Data Protection Directive (No. 95/46/EC)⁷ and the Convention No. ETS 108 of Council of Europe⁸. Fair management principles, established by Convention no. ETS 108 of the European Council and Directive 95/46/EC was transposed into LLPPD (Article 3). Legitimate criteria of processing General and special categories of personal data, stipulated in Articles 7 and 8, were included in the LLPPD Article 5. In addition the detailed provisions governing the powers of the supervisory authority in the field of data protection were included in the revision of this law⁹.

The new amendment to the law was enacted in April 13th 2004¹⁰. This amendment was adopted due to accession of the Republic of Lithuania to the European Union in May 2004. In the third version of the law there are some changes that have been adopted after experts' assessment of the Law on compliance with the norms of Directive 95/46/EC. The new version of LLPPD also includes the legal norms specific to Lithuania¹¹.

⁵ Law on Legal Protection of Personal data of the Republic of Lithuania, *Official Gazette*, 1996.06.11, Nr.: 63, Publ. Nr.: 1479

⁶ Law on Legal Protection of Personal data of the Republic of Lithuania, *Official Gazette*, 2000.07.31, Nr.: 64, Publ. Nr.: 1924

⁷ Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data. [interactive] [accessed 2013-05-03] <<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31995L0046:en:HTML>>

⁸ Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data. [interactive] [accessed 2013-05-02] <<http://conventions.coe.int/Treaty/en/Treaties/Html/108.htm>>

⁹ Asmens Duomenų Teisinės apsaugos Įstatymo Komentarai [The Comment of Personal Data Protection Law]. Vilnius, 2005. P. 4.

¹⁰ Law on Legal Protection of Personal data of the Republic of Lithuania, *Official Gazette*, 2004.04.24, Nr.: 60, Publ. Nr.: 2120

¹¹ Asmens Duomenų Teisinės apsaugos Įstatymo Komentarai. [The Comment of Personal Data Protection Law]. Vilnius, 2005. P. 4

Article 1 paragraph 2 of the LLPPD sets that this Law shall regulate relations arising in the course of the processing of personal data by automatic means, and during the processing of personal data by other than automatic means in filing systems: lists, card indexes, files, codes, etc. The Law shall establish the rights of natural persons as data subjects, the procedure for protecting these rights, the rights, duties and liability of legal and natural persons while processing personal data. The content of Article 1 paragraph 2 is transposed content of Article 3 paragraph 1 of the Directive. Processing by automatic means any personal data that is used in any IT system, regardless of how data mining is organized. Thus, the term includes, for example, and an unformatted text file, and structured data base. In addition, it is not important in what way a reference to certain data subject is done. Thus, this term includes not only text links (for example, name or identification code), but also audio and visual processing of personal data¹².

Article 1 paragraph 3 defines who falls within the scope of application of the law. Although with the respect to the Directive this part of LLPPD is redundant and on the scope of Article will not be commented further, but there is an evident gap in the law - is not considered a situation when the citizens of Lithuania are processing personal data in cloud computing resources, provided by third-party resources.

LLPPD article 1 paragraph 4 states that this Law shall not apply if personal data is processed by a natural person only for his personal needs not related to business or profession. Article 1 paragraph 4 corresponds to Directive Article 3 paragraph 2 section 2. LLPPD determine the scope of the law exceptions - that is to say, it does not apply where personal data are processed only for someone personal needs. This is in line with the Directive, but Directive also clearly identifies household activities. LLPPD, in turn, develops the concept of "personal needs" as an activity unrelated with business or profession¹³.

Article 2 of the Law defines the concepts that form the basis of the terms used in the Law. Paragraph 1 defines that personal data shall mean any information relating to a natural person (data subject) who is known or who can be identified directly or indirectly by reference to such data as a personal identification number or one or more factors specific to his physical, physiological, mental, economic, cultural or social identity. LLPPD definition of "personal data" corresponds to Article 2 of the Directive. A wide range of definition includes all information relating to a natural person, even the easily available and non-essential

¹² Asmens Duomenų Teisinės apsaugos Įstatymo Komentarai [The Comment of Personal Data Protection Law].. Vilnius, 2005. P. 13-15.

¹³ Idem, P. 17-18.

information and data. The term "information" includes audio, video, genetic data, fingerprints, etc. Opinion and subjective evaluations (eg, "embarrassed", "good customer", "Womanizer", "unreliable") - it is also within the meaning of the personal data defined by LLPPD. Identity of the person can be identified, directly or indirectly, from the data associated with any other information that the data controller has or can get. Personal identity can be determined by a person's name, code, image, genetic code, or a number of key criteria and factors group (address, date of birth, occupation, etc.). Even if the person has a nickname, his identity can be considered as one that can be determine if the controller knows the "key" code (or the list of nicknames), and can link with the appropriate entity. In this way we come to the definition of "relativity of identification." It can be illustrated by the following example: the name or surname and some additional information, related to them, are included in the database. From this additional information is impossible to set the identity of natural person. The name or surname is replaced by some nickname. A list of nicknames and their correspondence to names of real people is drawn up. If an organization has a database, and a list mentioned above, the information in possession of this organization must be considered as personal data. This organization could easily set identity of data subjects¹⁴.

Lithuanian national registry system has an accumulated huge amount of information about persons or objects of points, which is regulated by the State Information Resources Management Law¹⁵. It's data is one of the main sources of information to the authorities. State Information Resources Management Law provides some norms that regulate providing and usage of certain data to and by third parties. These norms are particularly important in regard to the processing of personal data. While personal information may be publicly available, the recipient is required to process the data only in accordance with the norms of LLPPD.

The remaining text of LLPPD regulates personal data controller functions, rights and duties and determine that the State Data Protection Inspectorate, accountable to the Government of the Lithuania, is responsible for the protection of Personal Data.

In summary it can be said that legal norms that regulates the protection of personal data in Lithuania (stricter than those of the European Union Directives) provides sufficient protection of personal data and severely limits the possibilities of uncontrolled processing of Personal Data. But they are lagging behind the speed of technological development.

¹⁴ Asmens Duomenų Teisinės apsaugos Įstatymo Komentarai. [The Comment of Personal Data Protection Law]. Vilnius, 2005. P. 24-25.

¹⁵ Law on State Information Resources Management of Lithuania, *Official Gazette*, Nr.: 163, Publ. Nr.: 7739

Technologies, such as cloud computing, reveal the gaps of LLPPD. In the following chapters it will be discussed how the current legal situation corresponds with the context of emerging needs of new personal data ecosystem and rapidly expanding usage of semantic technologies.

PERSONAL DATA ECOSYSTEM EVOLUTION TRENDS AND NEW APPROACHES

The world is changing fast. A new computing and information sharing architecture has emerged during the past 10 years. The policies, business models, social norms and technologies of today are simply different from what existed before. Analytics have become the new engine of economic and social value creation. The discovery and insights derived from linking previously disparate bits of data have become essential for innovation. More data is being collected, processed and transferred than ever before. Data is collected by billions of connected devices, people and sensors that record trillions of transactions and behaviors each day. The unprecedented amount of data being generated is created in multiple ways. Data is actively collected from individuals who provide it in traditional ways (by filling out forms, surveys, registrations and so on). They are also passively collected as a by-product of other activities (for example Web browsing, location information from phones and credit card purchases). The increasing use of machine-to-machine transactions, which do not involve human interaction, is generating significant amounts of data about individuals. All of this data is further analyzed and commingled to create inferred data¹⁶.

In addition, the definition of personal data is evolving. Traditionally, that definition was pre-determined and governed through the use of a binary approach: in most jurisdictions, the use of personally identifiable information (PII) was subject to strict restrictions whereas the use of non-PII was often uncontrolled. However, what is considered personal data is increasingly contextual; it changes with personal preferences, new applications, context of uses, and changes in cultural and social norms. Traditionally, organizations have used a variety of techniques to de-identify data and create value for society while protecting an individual's privacy. Such data was not subject to the same rules as PII, as an individual could not be identified from it. But technological advances and the ability to associate data across multiple sources is shifting boundaries of what is or is not PII, including potential re-identification of previously anonymous data. This issue is the subject of significant debate

¹⁶ World economic forum, *The World is Changing*. Unlocking the Value of Personal Data: From Collection to Usage. 2013. [interactive] [2013-05-04]
<http://www3.weforum.org/docs/WEF_IT_UnlockingValuePersonalData_CollectionUsage_Report_2013.pdf>
P. 7-8.

with some arguing that this means that all data is effectively personally identifiable and should be treated as such. Others urge caution, arguing that this would curtail many of the beneficial uses of anonymous data with minimal gains in privacy. A shift in approach to thinking less about the data and more about the usage could offer a way forward. If the usage impacts an individual directly it would require different levels of governance than data which is used in an aggregated and anonymous manner¹⁷.

The traditional data-protection approach, based on 1970s computing architectures in which governments and large organizations operated in discrete silos, was that the individual is involved in consenting to data use at the time of collection. The organization that collected the data then used it for a specified use, based on user consent, and then deleted the data when it was no longer needed for the specified purpose. That approach was appropriate when the data collection was often related to a specific service, a single organization or single use and when the computer data systems were not highly interconnected. Now, however, the walls of enterprise computing have opened up along with the data flows across traditional silos. Traditional approaches are no longer fit for the purposes for which they were designed, for several reasons:

- They fail to account for the possibility that new and beneficial uses for the data will be discovered, long after the time of collection.
- They do not account for networked data architectures that lower the cost of data collection, transfer and processing to nearly zero, and enable multiuser access to a single piece of data.
- The torrent of data being generated from and about data subjects imposes an undue cognitive burden on individual data subjects. Overwhelming them with notices is ultimately disempowering and ineffective in terms of protection – it would take the average person about 250 working hours every year, or about 30 full working days – to actually read the privacy policies of the websites they visit in a year.
- In many instances (for example, while driving a car or when data is collected using many machine to machine methods), it is no longer practical or effective to gain the consent of individuals using traditional approaches.
- Ensuring stakeholder accountability is a task that is increasingly challenging. Unlike the case 30 years ago, when the *big data* principles were established, the questions of

¹⁷ World Economic Forum, *Personal Data Ecosystem: Overview*. Personal Data: The Emergence of a New Asset Class. 2011. [interactive] [203.05.03]
<http://www3.weforum.org/docs/WEF_ITTC_PersonalDataNewAsset_Report_2011.pdf> P. 13-19.

“Who has data about you?” and “Where is the data about you located?” are impossible to answer today¹⁸.

The World Economic Forum’s and others (the Centre for Information Policy Leadership, Digital Advertising Alliance, the Asia-Pacific Economic Cooperation etc.) make efforts to convene a multiple stake holder dialogue, increase momentum to establish new and evolving norms to guide how personal data can be used to create value. The proposed European Commission Data Protection Regulation currently under discussion by the European Council and Parliament is the most comprehensive attempt to establish new norms for the flow of personal data. While differing views were expressed throughout the dialogue on these proposals and the underlying principles, it was clear that these rules when agreed will be the big step forward creating and implementing the new Personal Data ecosystem¹⁹.

However, issues of protection, security and the overall stewardship of personal data remain central to the global Personal Data ecosystem. While the complexity of operating in a decentralized and distributed networked environment poses new challenges, ensuring data security remains crucial.

PERSONAL DATA USAGE IN SEMANTIC TECHNOLOGIES

The new information technologies, in particular - semantic technologies, occupy important place in new Personal Data ecosystem. The rest of this article briefly describes the essence of semantic technologies and present the importance of Personal Data sets in machine reasoning and named entity recognition.

The term “*semantic technologies*” represents a family of technologies that have been in existence for a long time and seek to help derive meaning from information. Some examples of semantic technologies include natural language processing (NLP), reasoning, data mining, artificial intelligence (AI), category tagging, and semantic search. Many other modern technologies can be called semantic technologies. While all of these technologies have an overall goal in common - helping to make sense of large or complex sets of data without being supplied with any preordained knowledge about the data - they do not share much more

¹⁸ World Economic Forum, *Personal Data Covers All Aspects of Our Lives* Rethinking Personal Data: Strengthening Trust. 2012. [interactive] [2013-05-03] <http://www3.weforum.org/docs/WEF_IT_RethinkingPersonalData_Report_2012.pdf> P. 7.

¹⁹ World Economic Forum, *Personal Data Covers All Aspects of Our Lives* Rethinking Personal Data: Strengthening Trust. 2012. [interactive] [2013-05-03] <http://www3.weforum.org/docs/WEF_IT_RethinkingPersonalData_Report_2012.pdf> P. 34. Also Federal Trade Commission, *Personal Data Ecosystem*. Exploring privacy: a roundtable series. USA, 2010. [interactive] [2013-05-03] <<http://www.ftc.gov/bcp/workshops/privacyroundtables/personalDataEcosystem.pdf>>

than that. They are implemented using many different programming languages, produce data (signal) in many different formats, rely on very different underlying formalisms, and rarely work well together without investing a significant amount of effort in integration engineering.

Semantic Web technologies and *semantic technologies* both start with *semantic* is often a source of confusion. In brief, Semantic technologies are algorithms and solutions that bring structure and meaning to information. Semantic Web technologies specifically are those that adhere to a specific set of W3C (World Wide Web Consortium) open technology standards that are designed to simplify the implementation of not only semantic technology solutions, but other kinds of solutions as well²⁰.

Semantic technologies encodes meanings separately from data and content files, and separately from application code. This enables machines as well as people to understand, share and reason with them at execution time. With semantic technologies, adding, changing and implementing new relationships or interconnecting programs in a different way can be just as simple as changing the external model that these programs share²¹. With traditional information technology, on the other hand, meanings and relationships must be predefined and “hard wired” into data formats and the application program code at design time. This means that when something changes, previously unexchanged information needs to be exchanged, or two programs need to interoperate in a new way, the humans must get involved. Semantic technologies are “meaning-centered.” They include tools for auto recognition of topics and concepts, information and meaning extraction, and categorization. Given a question, semantic technologies can directly search topics, concepts, associations that span a vast number of sources.

Social Web sites, such as Facebook, YouTube, Delicious, Flickr and Wikipedia, and numerous other Web applications, such as Google and Amazon, rely on implicitly or explicitly collected data about their users and their activities to provide personalized content and services. As these applications become more and more connected on the Semantic Web, a major challenge is to allow various applications to exchange, reuse, and integrate user data from different sources. Such data comes in different flavors: user data such as user profiles, social networking/tagging/blogging data, etc. as well as usage data like click through data or query logs. The amount of people's data available on the Web is tremendously growing so that

²⁰ Weller K., *Knowledge Representation in the Social Semantic Web*. Berlin, New York: Walter de Gruyter, 2010. P. 117-123. P. 55.

²¹ Fill H.-G., Reischl I., *Stepwise Semantic Enrichment in Health-Related Public Management by Using Semantic Information Models*. Semantic technologies for business and information systems engineering: concepts and applications. Business Science Reference, 2012. P. 199-205.

sharing and mining these heterogeneous data corpora distributed on the Web is a non-trivial problem that poses several challenges to the Semantic Web community²².

Semantic interoperability between Social Web applications is becoming increasingly important as users leave a plethora of traces at diverse services on the Web. Semantic Web and Social Web technologies and paradigms provide means to facilitate integration of user and usage data, for example, with the principles of Linked Data and Microformats, vocabulary standards such as FOAF (friend-of-a-friend, used in social Web) and SIOC (semantically interlinked online community), standardized APIs (application programming interface) such as OpenSocial, or support for schema matching as provided by the Silk framework²³. Further, mechanisms like WebID, OpenId, OAuth (open standart for authorization) and FOAF+SSL allow for identification and authorization on the Social Web. Hence, the time is right to exploit and improve such technologies for connecting user and usage data traces on the Social Semantic Web.

Linking distributed traces of user data provides new possibilities for inferring and modeling user preferences and personalizing Web systems to individual needs. Novel models, techniques, frameworks and systems have to be developed to leverage Social Web semantics. While linkage of user and usage data promises advantages for recommendation and personalization, it also raises questions related to provenance, trust and privacy: how does one know that the data gathered from several sources can be trusted, and how can one avoid that sensitive personal data is disclosed to certain services or used to infer and expose sensitive information? Trust and privacy, and associated policies, may therefore impact mining and reasoning on the people's data.

Recognition of named entities (e.g. people, organizations, locations, etc.) is an essential task in many natural language processing applications nowadays. Named entity recognition (NER) is given much attention in the research community and considerable progress has been achieved in many domains, such as newswire or biomedical NER. Most NER systems use additional features like part-of-speech (POS) tags, shallow parsing, gazetteers, etc. Such kind of information requires external knowledge. A simple way to guess whether a particular phrase is a named entity or not is to look it up in a gazetteer or into Person Ontologies. The NER ontology is a set of mappings established manually between the taxonomies of named

²² Breslin J., Passant A., Decker S., *The Social Semantic Web*. Springer, 2009. P. 12-13.

²³ Breslin J., Passant A., Decker S., *The Social Semantic Web*. Springer, 2009. P. 1-3.

entity types²⁴. The traditional flat NER type sets consist of several general types (such as Organization, Person, Date, Location, Percent, Money). In the sentence „Dalia Grybauskaitė [class=Person] was in Kaunas [class=Location]” there are two named entities: Person and Location. Once named entity was identified and annotated, NER looks to external sources (ontology or gazetteer) to definition of named entity recognized. Extended Personal Data set collected in gazetteer or Person ontology is described in Table 1. The list can also include a nickname and the description of the pronunciation of person’s name, and much more.

Tab. 1. Attributes of Extended Named Entity recognition. Source:
http://nlp.cs.nyu.edu/ene/version7_1_0Beng.html

Attributes	Examples of Attribute Values	Freq	Typical ENEs of AV
Occupation	professional baseball player, economist, poet	46	Position_Vocation
Nationality	American, English, Swiss	29	Nationality, Country
Career	professor of experimental physics at Johns Hopkins University, Prime Minister of Canada	26	Position_Vocation
Work	Mona Lisa, Physical Optics, the United Nations Headquarters	25	Product, Facility
Alma Mater	Trinity College, Cambridge, Université Laval, Kyoto Imperial University	20	School
Place of Birth (City)	Paris, Manchester, Shanghai	19	City
Place of Birth (Province)	Massachusetts, Quebec, Sichuan	18	Province
Previous Stay	Singapore, New York, Tehran	12	Location
Teacher	Vladimir Nemirovich-Danchenko, Gaston Paris, Bernardo Pasquini	10	Person
Date of Death	August 11, 1955, 1778, unknown	10	Date
Era	the Edo period, the latter Kamakura period, the 11th century	8	Era
Award	the Nobel Prize the Golden Ball, Gold Gloves	8	Award
Real Name	François-Marie Arouet, Cecil Louis Troughton Smith, Liu Che	8	Person
Another Name	The Great Bambino, Shizong, Francisco	8	Person
Title	honorary doctorate by the University of Hamburg, knighted, elected to the Baseball Hall of Fame	6	Title
Competition	1998 FIFA World Cup, 1927 World Series, the Cardiff Singer of the World competition	6	Game
Father	George Herman Ruth, Sr., Thomas, Masukichi Kato	5	Person
Place of Death	New York, Birmingham, Passau	5	Location
Cause of Death	cancer, killed in a car crash, executed by decapitation	5	

In machine reasoning, the ontology provides us with knowledge related to the named entity type. In brief, ontology is a specification of a conceptualization. Ontologies are the

²⁴ Nadeau D., Sekine S., *A survey of named entity recognition and classification*. National Research Council Canada / New York University, 2007. [interactive] [2013.05.02] <<http://nlp.cs.nyu.edu/sekine/papers/li07.pdf>> P. 2-3.

structural frameworks for organizing information and they are used in artificial intelligence, the Semantic Web, systems engineering, software engineering, biomedical informatics, library science, enterprise bookmarking, information architecture, e-Governance²⁵, e-Learning, Police activity²⁶ etc. as a form of knowledge representation about the world or some part of it. In general, major uses of ontologies are human communication, system interoperability, system engineering²⁷. The Person Ontology is the specific kind of ontologies, which contain various data sets of Personal Data. For this reason, the ontology also is used at a first stage as a NER system²⁸.

With the Semantic Web under development, online privacy through access control is one serious issue that should be addressed²⁹. The Semantic Web offers an opportunity for access control policies to fully describe their access policies instead of resorting to usernames and groups, which rely on some pre-agreed contractual understanding that one individual or organization corresponds to some local username³⁰.

In Semantic technologies development Lithuania is 6-8 years behind of the European Union. The first semantic technologies projects in Lithuania are yet to be implemented. Only few universities and business enterprises are seriously involved in semantic technologies development. While scientific literature on this topic in Lithuanian has not been found, it can be said that semantic technologies developers in Lithuania have been facing or will face in the near future serious obstacles from the part of the Law on Legal Protection of Personal Data in his present revision. For this reason, there is urgent need of the reform of the Lithuanian legal

²⁵ It is still a long way for e-government to exploit the potentials of the Semantic Web. But it is reasonable to start preparing for it now and making use of the new technologies. On this subject see for example: Theocharis S., Tsihrintzis G., *Semantic Web Technologies in e - Government*. World Academy of Science, Engineering and Technology. Nr.64, 2012. P. 1238-1244. [interactive] [2013.05.05] <<http://www.waset.org/journals/waset/v64/v64-234.pdf>>

²⁶ For example, from December 2011 New Zealand Police have adopted locally developed semantic search technology in their three 24x7 contact Centers. The use of a customized dictionary of commonly used police acronyms, synonyms and operational terms allows communications Centre staff to search using their normal operational jargon. [<http://computerworld.co.nz/news.nsf/news/nz-police-deploy-semantic-search-technology>]

²⁷ Weller K., *Knowledge Representation in the Social Semantic Web*. Berlin, New York: Walter de Gruyter, 2010. P. 117-123.

²⁸ Nadeau D., Sekine S., *A survey of named entity recognition and classification*. National Research Council Canada / New York University, 2007. [interactive] [2013.05.02] < <http://nlp.cs.nyu.edu/sekine/papers/li07.pdf> >P. 4., Also Alexandros Valarakos*‡, Georgios Sigletos*, Vangelis Karkaletsis*, Georgios Paliouras* A Methodology for Semantically Annotating a Corpus Using a Domain Ontology and Machine Learning Department of Information and Telecommunication Systems Engineering, School of Sciences, University of the Aegean [interactive] [2013-05-04] <http://www.icsd.aegean.gr/incosys_old/publications/RANLP-NCSR-Valarakos.pdf> P. 2.

²⁹ Breslin J., Passant A., Decker S., *The Social Semantic Web*. Springer, 2009. P. 279-280.

³⁰ Shields B., Molloy O., Lyons G., Duggan J., *Securing Web Services Using Semantic Web Technologies Industrial Applications of Semantic Web Proceedings of the VIFIP WG72.5 Working Conference on Industrial Applications of Semantic Web, Finland*. Springer, 2005. P. 266.

framework of the definition and protection of personal data. At least, there is an urgent need for some intermediate solution on the protection of personal data. Otherwise, the right to privacy of the individual and public interest may suffer.

CONCLUSIONS

The legal norms that regulates the protection of personal data in Lithuania (stricter than those of the European Union Directives) provides sufficient protection of personal data and severely limits the possibilities of uncontrolled processing of Personal Data. But they are lagging behind the technological development speed. Technologies, such as cloud computing, reveal the gaps of LLPPD.

The World Economic Forum's and others (the Centre for Information Policy Leadership, Digital Advertising Alliance, the Asia-Pacific Economic Cooperation etc.) make efforts to convene a multiple stakeholder dialogue, increase momentum to establish new and evolving norms to guide how personal data can be used to create value. The proposed European Commission Data Protection Regulation currently under discussion by the European Council and Parliament is the most comprehensive attempt to establish new norms for the flow of personal data. While differing views were expressed throughout the dialogue on these proposals and the underlying principles, it was clear that these rules when agreed will be the big step forward creating and implementing the new Personal Data ecosystem.

The new information technologies, in particular - semantic technologies, occupies important place in new Personal Data ecosystem. In these technologies Personal Data are very important in NER, ontology creation and machine reasoning. In Semantic technologies development Lithuania stay behind of the European Union for 6-8 years. Their developers in Lithuania have faced or will face in the near future the serious obstacles from the part of the Law on Legal Protection of Personal Data in his present revision. For this reason, there is urgent need of the reform of the Lithuanian legal framework of the definition and protection of personal data. At least, there is an urgent need for some intermediate solution, waiting for the results of reform of the EU legal framework on the protection of personal data. Otherwise, the right to privacy of the individual and public interest may be infringed.

REFERENCES

1. Law on Legal Protection of Personal data, *Official Gazette*, 1996, Nr.: 63, Publ. Nr.: 1479.
2. Law on State Information Resources Management of Lithuania, *Official Gazette*, [2011](#) Nr.: 163, Publ. Nr.: 7739.

3. Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data. [interactive] [accessed 2013.05.02] <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31995L0046:en:HTML>
4. Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data. [interactive] [accessed 2013.05.02] <http://conventions.coe.int/Treaty/en/Treaties/Html/108.htm>
5. *Asmens Duomenų Teisinės apsaugos Įstatymo Komentarai*. Vilnius, 2005.
6. Biegel, S. *Beyond Our Control? Confronting the Limits of Our Legal System in the Age of Cyberspace*. The MIT Press, Cambridge, Massachusetts, 2001.
7. Breslin J., Passant A., Decker S., *The Social Semantic Web*. Springer, 2009.
8. Fill H.-G., Reischl I., *Stepwise Semantic Enrichment in Health-Related Public Management by Using Semantic Information Models*. Semantic technologies for business and information systems engineering: concepts and applications. Business Science Reference, 2012.
9. Nadeau D., Sekine S., *A survey of named entity recognition and classification*. National Research Council Canada / New York University, 2007. [interactive] [accessed 2013.05.02] <http://nlp.cs.nyu.edu/sekine/papers/li07.pdf>.
10. *NZ Police deploy semantic search technology*. Computerworld, 2012. [interactive] [accessed 2013.05.02] <http://computerworld.co.nz/news.nsf/news/nz-police-deploy-semantic-search-technology>.
11. Shields B., Molloy O., Lyons G., Duggan J., *Securing Web Services Using Semantic Web Technologies Industrial Applications of Semantic Web*. Proceedings of the VIFIP WG72.5 Working Conference on Industrial Applications of Semantic Web, Finland. Springer, 2005.
12. Theocharis S., Tsihrintzis G., *Semantic Web Technologies in e - Government*. World Academy of Science, Engineering and Technology. Nr.64, 2012. [interactive] [accessed 2013.05.02] <http://www.waset.org/journals/waset/v64/v64-234.pdf>
13. Valarakos A., Sigletos G., Karkaletsis V., Paliouras G., *A Methodology for Semantically Annotating a Corpus Using a Domain Ontology and Machine Learning*. Department of Information and Telecommunication Systems Engineering, School of Sciences, University of the Aegean [interactive] [accessed 2013.05.02] [http://www.icsd.aegean.gr/incosys_old/publications/RANLP-NCSR-Valarakos.pdf] P. 2
14. Weller K., *Knowledge Representation in the Social Semantic Web*. Berlin, New York: Walter de Gruyter, 2010.
15. World Economic Forum, *Personal Data: The Emergence of a New Asset Class* [interactive] [accessed 2013.05.02] http://www3.weforum.org/docs/WEF_ITTC_PersonalDataNewAsset_Report_2011.pdf
16. World Economic Forum, *Rethinking Personal Data: Strengthening Trust*. [interactive] [accessed 2013.05.02] http://www3.weforum.org/docs/WEF_IT_RethinkingPersonalData_Report_2012.pdf
17. World Economic Forum, *Unlocking the Value of Personal Data: From Collection to Usage* [interactive] [accessed 2013.05.02] http://www3.weforum.org/docs/WEF_IT_UnlockingValuePersonalData_CollectionUsage_Report_2013.pdf
18. Federal Trade Commission, *Personal Data Ecosystem*. Exploring privacy: a roundtable series. USA, 2010. [interactive] [accessed 2013.05.02] <http://www.ftc.gov/bcp/workshops/privacyroundtables/personalDataEcosystem.pdf>.

ASMENS DUOMENŲ NAUDOJIMAS SEMANTINIŲ TECHNOLOGIJŲ VYSTYMO KONTEKSTE: TEISINIAI KLAUSIMAI

Darius Amilevičius*

Mykolo Romerio universitetas

Birutė Pranevičienė**

Mykolo Romerio universitetas

Santrauka

Skaitmeninės revoliucijos, spartaus internetinių technologijų plėtros, milžiniškais tempais didėjančios informacijos ir duomenų apimties, virtualių socialinių tinklų ekspansijos kontekste asmens duomenų apsaugai kyla vis nauji iššūkiai. Atlikus teisės aktų ir informacinių technologijų plėtros tendencijų analizę, straipsnyje konstatuojama, kad globalizacija ir naujos asmens duomenų ekosistemos koncepcija verčia persvarstyti tradicines asmens duomenų ir jų apsaugos sampratas bei jų santykį su viešuoju interesu, kurį atstovauja skaitmeninės ekonomikos, mokslo ir technologijų plėtra. 2012 m. Europos Komisija pasiūlė iš esmės reformuoti Europos Sąjungos asmens duomenų apsaugos teisės aktus. Siūloma reforma turėtų sustiprinti individualias teises, spręsti globalizacijos keliamas problemas bei naujas technologijų plėtros iššūkius. Straipsnyje atskleidžiama, kad aktuali Lietuvos Respublikos Asmens duomenų teisinės apsaugos įstatymo redakcija, griežčiau už Europos Sąjungos teisinės normas reguliuojanti asmens duomenų naudojimą ir jų apsaugą, apsunkina pažangių kalbos ir semantinių technologijų plėtrą Lietuvoje, nors šioje srityje Lietuva 6-8 metais atsilieka nuo Europos Sąjungos vidurkio. Autoriai prieina išvados, kad susidariusioje situacijoje, nelaukiant Europos Sąjungos asmens duomenų apsaugos reformos rezultatų, Lietuvoje būtina tarpinė asmens duomenų apsaugą reguliuojančių teisės aktų reforma, nes susidariusi teisinė ir technologinė situacija kelia rimtą grėsmę asmens teisei į privatumą, asmens duomenų apsaugai bei viešajam interesui.

Pagrindinės sąvokos: asmens duomenų apsauga, semantinės technologijos, teisė į privatumą.

Darius Amilevičius*, Mykolo Romerio universiteto Viešojo saugumo fakulteto Humanitarinių mokslų katedros docentas. Mokslinių tyrimų kryptys: elektroninės erdvės saugumo ekonomika, kompiuterinė lingvistika, politinė ir juridinė retorika, žmogaus teisės.

Darius Amilevičius*, Mykolas Romeris University, Faculty of Public security, Department of Humanities, assoc. professor. Research interests: cybersecurity economics, computational linguistic, political and juridical rhetoric, human rights.

Birutė Pranevičienė**, Mykolo Romerio universiteto Viešojo saugumo fakulteto Teisės katedros profesorė. Mokslinių tyrimų kryptys: konstitucinė teisė, žmogaus teisės, aplinkos teisė.

Birutė Pranevičienė**, Mykolas Romeris University, Faculty of Public security, Department of Law, professor. Research interests: constitutional law, human rights, environmental law