Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 104 (2017) 57 – 64

Procedia Computer Science

ICTE 2016, December 2016, Riga, Latvia

Service-Oriented Architecture of Intelligent Environment for Historical Records Studies

Aleksandrs Ivanovs[a,b], Aleksey Varfolomeyev[c,*]

[a]*Daugavpils University, Vienības iela 13, Daugavpils, LV-5401, Latvia*
[b]*Institute for Regional Studies (REGI), Rezekne Academy of Technologies, Atbrīvošanas aleja 115, Rēzekne, LV-4601, Latvia*
[c]*Petrozavodsk State University, Lenina pr. 33, Petrozavodsk, 185910, Russia*

Abstract

Nowadays, such socio-technical systems as virtual research environments are increasingly employed in the studies of textual heritage. The paper presents a modular platform based on a combination of Web services that provide tools for historical records publishing and research. A promising approach is the creation of semantic networks that encompass marked up texts of records (in this case, medieval documents preserved at Latvian State Historical Archives) and corresponding meta-information. The authors describe the prototype of a software system, which maps XML markup into ACE (Attempto Controlled English) statements; these statements can be automatically translated into OWL (Ontological Web Language), visualized as semantic networks, or queried by means of Web services of Attempto Project. This system is designed as an intelligent environment for collaborative research.

Peer-review under responsibility of organizing committee of the scientific committee of the international conference; ICTE 2016

Keywords: Virtual research environment; Service-oriented architecture; Digital humanities; Attempto controlled English

1. Introduction

Rapid development of computer networks technologies has resulted in their introduction into almost all spheres of life, including scholarly research. Within this trend, virtual research environments[1] (VRE) are widely employed by distributed communities of researchers in order to organize collaborative processing of distributed data. At present,

* Corresponding author. Tel.: +7-953-536-0268; fax: +7-814-2711000.
E-mail address: avarf@petrsu.ru

Peer-review under responsibility of organizing committee of the scientific committee of the international conference; ICTE 2016
doi:10.1016/j.procs.2017.01.062

VRE have developed into multiplex socio-technical systems that embrace common software environments, which can be adapted for different research tasks, software components, which are used as research tools, as well as users who form research teams[2]. In this regard, the problem of interaction between the constituent parts of such systems is of utmost importance. For instance, it is desirable to ensure an exchange of research methods (algorithms) between distributed users. Moreover, the set of research tools elaborated for a definite project might be supplemented with the tools derived from other research environments.

Nowadays, many research tools are employed as autonomous software agents that interact with other programs by means of API. A common approach to the elaboration of such agents is their design in the form of Web services[3,4] (SOAP-based or RESTful). Interfaces of Web services are described using standardized languages; it makes possible to apply to these Web services from other programs.

A number of specialized Web services have been created for VRE in the field of Digital Humanities[5,6]. For example, in CLARIN Project[7] that unites national language resources, dozens of Web services aimed at processing of the texts in different European languages are used. Most of the problems that can be settled by means of these Web-services are also topical in historical records studies and history research in general (e.g., named entity recognition, event extraction, sentiment analysis, etc.). However, in historical records studies, the scope of information that should be extracted from historical records is much wider; therefore, the aforementioned Web services are quite insufficient to process original historical records.

One of the noticeable trends in historical records studies is semantic representation of historical information and, in particular, the sense of the texts of historical sources, which is recorded in a formalized way suitable for computer processing[8,9,10]. Actually, this trend reflects the topicality of numerous projects aimed at representation of cultural heritage; moreover, it corresponds fully with the Semantic Web paradigm[11].

In compliance with this trend, the authors have designed a modular platform, which is based on a combination of Web services that provide tools for semantic publishing and research of historical records. A promising approach is the creation of semantic networks that encompass marked up texts of records (in this case, medieval documents[12] preserved at Latvian State Historical Archives – a structural unit of the National Archives of Latvia) and corresponding meta-information. In the paper, the authors describe the prototype of a software system, which maps XML markup into ACE statements; these statements can be automatically translated into OWL, visualized as semantic networks, or queried by means of Web services of Attempto Project[13]. The system is designed as an intelligent environment for collaborative research.

The rest of the paper is organized as follows. Section 2 provides a brief overview of Attempto Controlled English and Web services specially developed for this controlled language. Section 3 presents the architecture of a virtual research environment designed by the authors for the purposes of an in-depth study of historical records, which are published in XML-format. Section 4 describes an example of application of the virtual research environment to the study of medieval charters in Old Russian. Section 5 shows how this software environment can be used to supplement and enrich semantic resources aggregated for historical e-tourism and museums. Section 6 summarizes the results of the paper.

2. Attempto controlled English

Attempto Controlled English (ACE) is one of the controlled natural languages[14] that are used for knowledge representation. ACE is a subset of Standard English with a restricted syntax. It means that ACE is understandable by both humans and computers.

ACE texts can be translated into the so-called Discourse Representation Structures[15] (DRS) that, in their turn, can be directly correlated with first-order logic formulas. Therefore DRS can be translated into any formal language (e.g., Rule Markup Language, Semantic Web Rule Language, OWL, etc.), which is equivalent to a subset of first-order logic. Such representation of the texts provides a basis for logic reasoning, namely: automatic question answering, theorem proving, and logical entailment.

It is very important that in the course of processing ACE texts different domain-specific lexicons can be used. Furthermore, ACE-texts can be processed using reasoner, which generate new hypotheses on the bases of the facts revealed by a researcher. Therefore, this language is quite appropriate for the purposes of semantic publication of historical records.

2.1. ACE grammar

The grammar of Attempto Controlled English defines and constrains the form and the meaning of ACE sentences and texts. ACE's grammar is expressed as a small set of construction rules, and the meaning of ACE texts is defined by a small set of interpretation rules. These rules simplify the Standard English language. Meantime, ACE grammar and construction rules are rather flexible.

The vocabulary of ACE is restricted to the set of predefined words and some predefined phrases. However, this set can be extended by a user. The verbs can be used in the Present Indefinite Tense only; however, there can be used both the Active and the Passive Voices. Each noun should be introduced by an article or a determiner (a, every, no, some, etc.), with the exception of proper names, which are written with capital letters.

It is important, that in ACE not only simple sentences, but also composite and interrogative sentences can be constructed. If-then sentences are also used to specify conditional statements.

Although this language puts some substantial restrictions on text rendering, nevertheless, it seems that ACE construction rules can be applied to the texts of historical records, including medieval charters. The translation of the texts into ACE comprises, to a certain extent, researchers' interpretations of the sense of documents.

2.2. ACE Web services

Attempto Controlled English is supported by a number of software tools. Within Attempto Project, three of them – Attempto Parsing Engine, ACE Reasoner, and AceRules– have been designed as Web services.

Attempto Parsing Engine[13] (APE) is a parser that translates ACE texts into formal logic languages (DRS, OWL, etc). APE is designed as a RESTful Web service. It can be accessed through HTTP: ACE-texts and queries are fed in, and, according to the parameters of initial queries, answers are formulated in different formal languages. One of the output formats is OWL in XML syntax that provides a plenty of means of processing and visualization in the form of a semantic network.

ACE Reasoner[16] (RACE) is a SOAP-based Web service which runs in three modes, namely, as a checker of consistency of a set of axioms, as a theorem prover, and, as a question answering system. For instance, in query answering mode, a set of ACE-sentences (axioms) and one question formulated in ACE as an interrogative sentence are inputted. As an output, APE provides nouns and proper names, which are relevant to the query, as well as a list of axioms that have been used to prepare the answer. Both queries and answers are designed as the so-called SOAP-envelopes by means of special XML based language.

AceRules[17] is a SOAP-based Web service, which provides an opportunity to draw all possible conclusions from a given set of ACE statements. Unlike RACE, this Web service allows employing users' dictionaries (URL of a dictionary can be indicated). Moreover, as in the case of RACE, an input should contain a number of axioms recorded in ACE, as well as parameters that determine an output. Answers are formulated in ACE as sets of possible inferences from axioms.

3. The architecture of historical records studies environment

As it is shown in the previous Section, ACE Web services are promising tools that provide for extraction of semantic information from ACE texts. If the texts of historical documents are written in ACE, a researcher will have additional opportunities to create semantic publications of historical records, as well query answering systems and tools for automatic advancement and verification of hypotheses. We propose a prototype of a Historical Records Studies Environment (HRSE) that is based on ACE Web services. It should be noted that a prerequisite for the correct functioning of this system is XML-markup of historical documents[18,19], which are fed in HRSE.

3.1. Mapping XML to ACE

In recent decades, such XML-markup schemes as TEI (www.tei-c.org) and CEI (www.cei.lmu.de) have been increasingly used for the production of electronic publications of historical sources. Semantically arranged historical information can be extracted from XML-publications of historical documents, since TEI and, especially, CEI make

it possible to mark up diverse historical objects in the texts of the documents. For this purpose, a number of CEI tags <persName>, <roleName>, <issuer>, and <addressee>, as well as a universal tag <name> in combination with the attribute 'type', are used. To create a knowledge base on the basis of XML-markup, a program that either directly translates XML into OWL, or generates statements in ACE from XML-texts (i.e., CEI-markup) should be written.

We have designed a special Web application XML2ACE for HRSE. A user of the application feeds in the rules of XML mapping to ACE, which are formulated as sets of pairs (XML pattern and ACE-statement). The output is an ACE-text. Table 1 shows an example of rules for CEI mapping to ACE.

Table 1. An example of rules for mapping CEI to ACE.

CEI pattern	ACE statement
<charter id="X">	X is a charter.
<charter id="X">…<issuer reg="Y"> …</issuer>…</charter>	Y sends X. X mentions Y.
<roleName reg="role_x">	Every role_x is a person.
<issuer reg="Y">…<roleName reg="role_x">…</roleName>…</issuer>	Y is a role_x.
<charter id="X">…<roleName reg="role_x">…</roleName>…</charter>	X mentions a role_x.
<name type="city">Z</name>	Z is a city.
<issuer reg="Y">…<name type="city">Z</name>…</issuer>	Y lives in Z.
<charter id="X">…<name type="city">Z</name>…</charter>	X mentions Z.

In HRSE, automatically generated ACE-statements can be supplemented with axioms formulated as if-then statements that are related to a definite domain of knowledge. After that, the ACE texts can be processed applying such Web services as APE, RACE, or AceRules.

3.2. Web services cooperation

An ACE-text generated from XML-markup of a historical record forms a specific knowledge base that represents extensive information about this document. For this purpose, one should employ definite means that allow extracting knowledge from the texts. As a matter of fact, ACE Web services are able to perform some tasks related to knowledge extraction. For instance, applying Webservice APE, on the basis of ACE-texts, OWL ontologies can be created. On the Semantic Web, such ontologies are generally used to represent knowledge within a definite domain. Moreover, there are many programs used for visualization of ontologies in the form of semantic networks. One of such programs is OWLGrEd[20] elaborated at the University of Latvia that makes the process of visualization easier and the results of visualization very convincing.

In historical records studies, ACE-texts can be also used to make logical inference. The tools of logical inference make it possible to advance and verify hypotheses, to put semantic queries to the texts, to run a computer in a query answering mode, to draw a number of presumable inferences on the basis of a set of premises. For this purpose, Web services RACE and AceRules are used. It stands to reason that in order to draw a logical inference, a user has to work out a set of additional inference rules (axioms) correlated with a definite domain of knowledge.

In order to take the advantages that might be provided by the system, HRSE is designed as a Web application that within one and indivisible environment offers such opportunities as mapping XML to ACE, editing ACE-texts simultaneously preserving their integrity, applying appropriate Web services that have diverse initial data, and preserving the results of ongoing research. The workflow based on Web services cooperation in the Historical Records Studies Environment is shown in Fig. 1.

4. Case study

The empirical basis of this paper embraces Old Russian charters dating back to the 13th century. These historical records , which form a constituent part of the vast collection of medieval and early modern records "Moscowitica–

Ruthenica"[12] kept at the Latvian State Historical Archives (since 2011, a structural unit of the Latvian National Archives of Latvia); record groups No. 8 and 673.

Let us consider five interconnected charters, which reflect the course of relations between Riga and Smolensk in the 13th century. The charters have been marked up according to the CEI-markup scheme; after that the texts have been translated into ACE, employing XML2ACE described in the Section 3.1.

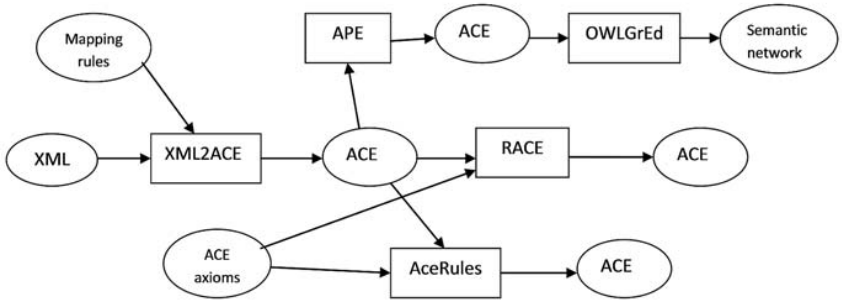

Fig. 1. Web services cooperation in historical records studies environment.

First of all, we have built a semantic network that represents meta-information, i.e., facts about the aforementioned five charters. Previously, a similar semantic network was built manually on the basis of expert evaluations of the texts' tenor; this variant of the semantic network was described in research papers. To build the network automatically, the Web service APE has been employed to map ACE-text of the charters to OWL in XML syntax. Then, the obtained OWL-text has been fed in OWLGrEd Web client, which, in its turn, has generated the semantic network (see Fig. 2). It should be noted that this image of the semantic network is more informative than the manually drawn visualization of the same network.

However, more difficult research questions can be posed. In the charters, three persons are mentioned as governors of Smolensk: Artemii in charter No. 4, Grigorii in charter No. 5, and an unknown person who in charter No. 6 bears the identifier "C6p3". Taking into consideration the fact that the charters were compiled during one and the same year, we can assume that either the governors were replaced, or there was made a mistake in the text of the document. Anyway, we have enough data to conclude that the unknown person is either Artemii or Grigorii. The researcher's logic that allows reaching the aforementioned conclusion can be recorded as the following axiom (rule): "If X is a governor and X lives in Y and Z is a governor and Z lives in Y then X equals Y likely."

We have made an attempt to apply RACE and AceRules to formulate the hypothesis that the unknown person might be identified either as Artemii or Grigorii. The question that was inputted in RACE was: "Who equals C6p3 likely?" RACE failed to give an answer to this question; at the same time, it was able to answer to more simple questions (e.g., "Who is mentioned by Charter6?") [21]. Unlike RACE, AceRules gave an answer to the first question. Actually, it even provided a plenty of conclusions, including: "C6p3 equals Artemii likely"; "C6p3 equals Grigorii likely". It is obvious that a researcher can also make such suppositions. Unfortunately, there were a number of banal statements and conclusions, such as "Grigorii equals Grigorii". It is clear that these statements are quite correct. In perspective, applying text processing services, banal statements might be detached from informative ones.

5. Additional possibilities

Historical Records Studies Environment designed by the authors of the paper can be also used as a constituent part of other socio-technical systems that process semantically enriched information. As a rule, in such systems, just elaboration of semantic structures (e.g., RDF-triplets) on the basis of unstructured data requires maximum resources.

Moreover, socio-technical systems oriented to wide ranges of users should have convenient human-computer interfaces to put queries to knowledge bases. Some additional possibilities provided by HRSE have been revealed in the two projects aimed at elaboration of intelligent environments on the basis of semantic space approach[22].

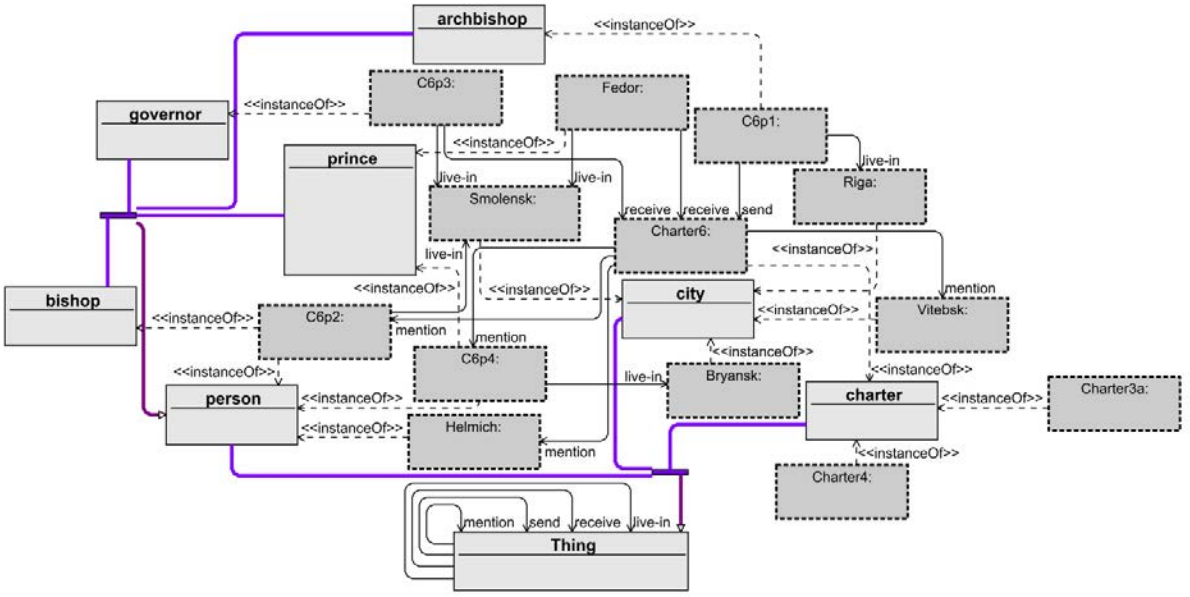

Fig. 2. Image of the semantic network powered by OWLGrEd.

5.1. Recommender service for historical tourism

The architecture of a recommender service that accompanies those tourists who want to receive historical data related to definite points of interest (POIs) has been described in a number of papers[23,24,25,26]. The service employs software agents, which retrieve data from available sources, build a semantic network around a definite POI on the basis of the collected information, and provides recommendations to visit other POIs that are interconnected with the initial POI by means of common or related historical events and/or persons.

The deciding factor that ensures proper functioning of the recommender service is existence (and availability) of information resources suitable for the network building. Unfortunately, at present, historical information is underrepresented on the Semantic Web. At the same time, rather a lot of historical sources (documents, narratives, etc.) have been published in the form of XML-files according to TEI and CEI markup schemes (e.g., see project monasterium.net, etc.). Thus, it is quite topical to map such XML-documents to Semantic Web formats (OWL). For this purpose, the tools of HRSE described in Section 3 might be employed; and the workflow might embrace three stages:

- For each collection of XML-documents, rules for mapping XML to ACE are created manually
- Applying XML2ACE, XML-texts are mapped to ACE-texts
- Employing APE, the ACE-texts are mapped to OWL and processed, using software agents of the recommender service

It should be noted that this approach to the production of semantic data resources for the recommender service based on such lightweight tools as ACE or Semantic Wiki has been already discussed[25,26]. However, up to recent times, appropriate tools for the implementation of the approach have not been developed.

5.2. Museum information service

Recommender services like that described in Section 5.1 can be developed for other purposes also. For instance, a semantic service intended for employees and visitors of museums can be created[27]. On the basis of information about exhibits, this service should build a semantic network focused on definite museum objects and, thus, bring other related objects, which are represented within the semantic network, to visitors' and/or museum employees' attention. Unlike the example discussed in the previous section, in this case, information sources represented as XML-files are not used (the only exception might be the links to external collections of marked up historical records). Yet, there emerges a problem of user-friendly semantic annotations that should be made on museum objects. It is desirable that the annotations should be made in ACE by museum employees.

These annotations can be fed in HRSE and then, employing Web service APE, mapped to OWL. It should be noted that ACE-texts as such can be directly used to inform museum visitors. Moreover, Web services RACE and AceRules can also be used to establish a dialogue with visitors: in accordance with inference rules, which are formulated by museum employees, answers to visitors' questions can be offered and diverse inferences from facts can be drawn.

6. Conclusion

In the paper, the prototype of VRE (HRSE) designed for an in-depth research and representation of collections of historical records, which are marked up according to XML markup schemes, is described. Within HRSE, Web services recently elaborated in ACE Project as well as a new Web service that maps XML to ACE are used. In the course of testing the prototype, medieval documents preserved at Latvian State Historical Archives (Riga) have been used. The forthcoming research will be focused on evaluation of the efficiency of the system as well as on productivity increasing and creation of a user-friendly interface.

Acknowledgements

This work is financially supported by the Ministry of Education and Science of the Russian Federation within project # 2.2336.2014/K from the project part of state research assignment, by the project # 14.574.21.0060 (RFMEFI57414X0060) of Federal Target Program "Research and development on priority directions of scientific technological complex of Russia for 2014–2020", and by Russian Foundation for Humanities according to research project # 16-01-12033.

References

1. Candela L et al. Virtual Research Environments: An Overview and a Research Agenda. *Data Science Journal*. 12; 2013. p. 75–81.
2. de Roure D, Goble C, Stevens R. The design and realisation of the myExperiment Virtual Research Environment for social sharing of workflows. *Future Generation Computer Systems*. 25; 2009. p. 561-567.
3. Singh M, Huhns M. *Service-Oriented Computing*. Wiley; 2005.
4. Blanke T, Hedges M, Palmer R. Restful Services for the e-Humanities – Web Services that Work for the e-Humanities Ecosystem. In: 3rd IEEE International Conference on Digital Ecosystems and Technologies. IEEE; 2009. p. 637–642.
5. Aschenbrenner A, Gietz P, Kuster MW, Ludwig Ch, Neuroth H. TextGrid – a Modular Platform for Collaborative Textual Editing. In: *Proceedings of the International Workshop on Digital Library Goes e-Science (DLSci06)*. Alicante; 2006. p. 27–36.
6. Blanke T et al. Deploying general-purpose virtual research environments for humanities research. *Phil. Trans. R. Soc;* 2010. p. 3813-3828.
7. Hinrichs E, Krauwer S. The CLARIN Research Infrastructure: Resources and Tools for E-Humanities Scholars. In: *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC-2014)*. p.1525–1531.
8. Ahonen E, Hyvonen E. Publishing Historical Texts on the Semantic Web – A Case Study. In: *Proceedings of the Third IEEE International Conference on Semantic Computing (ICSC2009)*. Berkeley (CA); 2009. p. 167–173.
9. Ivanovs A, Varfolomeyev A. Some approaches to the semantic publication of charter corpora. Böhlau Verlag. Köln, Weimar, Wien; 2014. p.149-168.
10. Merono-Penuela A et al. Semantic technologies for historical research: A survey. *Semantic Web journal*. 6 (6); 2015. p. 539–564.
11. Shadbolt N, Berners-Lee T, Hall W. The Semantic Web Revisited. *IEEE Intelligent Systems*. 21 (3); 2006. p. 96–101.

12. Ivanovs A, Kuznecovs A. *Smolenskas–Rigas aktis: 13.gs.–14.gs. pirma puse: Kompleksa Moscowitica – Ruthenica dokumenti par Smolenskas un Rigas attiecibam*. Riga; 2009. 768. (in Latvian).
13. de Coi JL, Fuchs NE, Kaljurand K, Kuhn T. Controlled English for Reasoning on the Semantic Web. In: Bry F, Maluszynski J. (eds.) Semantic Techniques for the Web: The Reverse Perspective. Heidelberg: Springer-Verlag; 2009. p. 276–308.
14. Kuhn T. A Survey and Classification of Controlled Natural Languages. *Computational Linguistics*. 40 (1); 2014. p. 121-170.
15. Kamp H, Reyle U. *From Discourse to Logic: Introduction to Modeltheoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory*. Dordrecht: Kluwer Academic Publishers; 1993. 716.
16. Fuchs NE. First-Order Reasoning for Attempto Controlled English. In: *Proceedings of the Second International Workshop on Controlled Natural Language (CNL 2010)*. Springer; 2012. p. 73-94.
17. Kuhn T. AceRules: Executing Rules in Controlled Natural Language. In: Marchiori M, Pan JZ, de Sainte CM. (eds.). *Proceedings of the First International Conference on Web Reasoning and Rule Systems (RR2007)*. Springer; 2007. p.299-308.
18. Ivanovs A, Varfolomeyev A. Editing and Exploratory Analysis of Medieval Documents by Means of XML Technologies: The Case of the Documentary Source Complex Moscowitica–Ruthenica. In: *Humanities, Computers and Cultural Heritage: Proceedings of the Association for History and Computing*. Amsterdam: Royal Netherlands Academy of Arts and Sciences; 2005. p. 155–160.
19. Burkard B, Vogeler G, Gruner S. Informatics for Historians: Tools for Medieval Document XML Markup, and their Impact on the History-Sciences. *Journal of Universal Computer Science*. 14 (2); 2007. p. 193-210.
20. Barzdins J et al. UML Style Graphical Notation and Editor for OWL 2. In: Perspectives in Business Informatics Research. Lecture Notes in Business Information Processing. Vol. 64 (2); 2010. p. 102-114.
21. Varfolomeyev A, Ivanovs A. Representation of Historical Sources on the Semantic Web by Means of Attempto Controlled English. *Communications in Computer and Information Science*. Vol. 394. Springer; 2013. p. 177-190.
22. Korzun D, Balandin S, Gurtov A. Deployment of Smart Spaces in Internet of Things: Overview of the design challenges. In: Balandin S, Andreev S, Koucheryavy Y. (eds.). Proc. 13th Int'l Conf. Next Generation Wired/Wireless Networking and 6th Conf. on Internet of Things and Smart Spaces (NEW2AN/ruSMART 2013). LNCS 8121. Springer; 2013. p. 48–59.
23. Varfolomeyev A et al. Smart Spaces Approach to Development of Recommendation Services for Historical e-Tourism. In: UBICOMM 2015. IARIA; 2015. p. 56-61.
24. Varfolomeyev A et al. Smart Space based Recommendation Service for Historical Tourism. *Procedia Computer Science*. Vol. 77; 2015. p. 85–91.
25. Varfolomeyev A et al. Smart space services as a teaching aid in history and cultural heritage studies. In: *ICERI2015 Proceedings*. 8th International Conference of Education, Research and Innovation. IATED; 2015. p.1373-1380.
26. Varfolomeyev A, Korzun D, Ivanovs A, Petrina O. Smart Personal Assistant for Historical Tourism. In: Arapatsakos C, Razeghi M, Gekas V. (eds.). *Proc of the 2nd Int'l Conf on Environment, Energy, Ecosystems and Development (EEEAD 2014)*; 2014. p. 9-15.
27. Marchenkov S, Vdovenko A, Petrina O, Korzun D. A Smart Space-Based Design of Semantic Layer for Advancing Museum Information Services. In: *Proc 19th Conf of Open Innovations Association FRUCT*. Yvaskyla Univeristy; 2016.

Aleksandrs Ivanovs graduated Moscow State Historical Archival Institute, Faculty of Historical Archives. Dr. hist. (1993, University of Latvia). Professor in Historiography and Auxiliary Historical Disciplines (2007). Present positions: Professor (Daugavpils University, Faculty of Humanities, History Department) and Leading Researcher (Rezekne Academy of Technologies, Institute for Regional Studies – REGI). Member of the Commission of the Historians of Latvia. About 150 publications – research paper and monographs in the historiography of Latvia, source studies, historical diplomatics, and computer-based historical source studies. Contact him at aleksandrs.ivanovs@du.lv.

Aleksey Varfolomeyev is an associate professor at Petrozavodsk State University (Department of Theory of Probability and Data Analysis). He graduated from the Faculty of Mathematics at Omsk State University in 1983. He received PhD degree in Mathematical Cybernetics at Saint-Petersburg State University (Russia) in 1994. His research interests include mathematical models of text structure and content, Semantic Web technologies for digital humanities. Contact him at avarf@petrsu.ru.