An Ontology Approach to Storing Educational Information

Aleksandra Vatian, Natalia Gusarova, Galina Artemova, Natalia Dobrenko
ITMO University
Saint Petersburg,, Russian Federation
alexvatyan@gmail.com, natfed@list.ru, g.o.artemova@gmail.com, graziokisa@ya.ru

***Abstract*—The modern educational process is characterized by the fact that the competencies and the content of education quickly change and dynamically vary depending on the needs of the specific groups of students. Especially strongly it is shown in the professional education, for example in information technologies. The teacher should have an opportunity to quickly and effectively arrange the content of education for the parameters of a specific course (for instance: type of group, periods of training, etc.). In these conditions, the traditional method of storing the information, such as the relational data model, seems difficult and inconvenient to manage with the permanently and dynamically changing structure of data. In this regard, the use of network models of data structure is required. In this paper, we present an ontological model for an information system for storing educational information. The authors propose the use of a wiki-engine for building an information system and the SCORM-based organization of the training material. The experimental results have shown that the ontological approach for storing the educational information tends to cover the needs of the teachers in universities.**

I. INTRODUCTION

The modern educational technologies are based on the application of information technologies, automation equipment and electronic information resources. At the same time special attention is paid to the development of technologies used for automating the educational processes. The amount of learning material, needed to support the educational process, has grown rapidly in the recent years. As a result, the teachers in universities are faced with the challenge of storing and managing the educational material quickly and effectively. The amount of material managed by the teachers is really huge. The educational programs, that they need to manage, determine the educational content of a certain level and directivity. In particular, programs of higher education include programs of a bachelor and magister degree, take into account the educational needs and the requests of the students, include the curriculum, working programs of training courses, objects, disciplines (modules) and other materials.

The working program of a discipline is developed by the leading it teacher and includes the list of documents based on which it is developed, the purposes and tasks of discipline, the volume and the content of discipline, all the methodical support of the discipline (the main and additional literature, technical and audiovisual training aids, application program packages, materials for the practical works, homework, methodical instructions to students). Moreover, the teachers have to permanently update the available material, to delete the irrelevant parts and to adapt it for the new courses.

Despite a great deal of research in the area of e-learning and distance learning systems, little work has been done for creating informational systems for storing educational materials for the teachers. In other words, current works and standards are focused on systems that are created for static material and do not assume dynamic changes and reformatting to specific course requirements. The purpose of this research was to design such system that can be used by the teachers and that would be enough convenient for storing editing and using the teaching material. To achieve this goal, we propose an approach based on ontologies to represent the educational material.

This paper is organized as follows: Section 2 presents the related work according to the research issues; Section 3 describes the architecture of the information system, while Section 4 presents the testing experiments. Finally, in Section 5 the conclusion and future work are outlined.

II. RELATED WORK

Due to the rapid growth of information in various fields the problem of efficient information support of scientific and teaching activities, as well as decision-making processes becomes more and more urgent.

Typically, this problem is considered in the context of creating the storage of documents and their systematization in order to facilitate the search for the necessary information.

However, there are not enough opportunities given by this approach to effectively support the scientific educational activities. Data is mainly represented in existing systems as text documents or relational databases, and the interpretation and the data representation as the

ISBN 978-952-68397-6-9 (paperback), ISBN 978-952-68397-7-6 (PDF)

connected parts of the material, as a rule, are not supported and rely on the user. At the same time, for the teacher, the most needful form of information representation is to present it in the form of a set of many interrelated facts or documents.

So it is very important to have the ability to analyze the documents, for example, to compare them and / or to determine rapidly the source of this document, to find the set of facts associated with each other and etc. That is why it seems that the only opportunity to manage these requests is to provide the information system that uses the semantic ontology data model. Due to this fact, there is a problem of developing such information systems, where the information is represented as a knowledge network data.

In this regard a large number of researches represent educational content with the use of ontological approach, but most of them, as it was mentioned earlier, are aimed to be implemented as the e-learning and distance learning systems by students. Many systems were proposed to manage the documents effectively [8].

In [7] the authors propose a Semantic Ontology Mapping for Interoperability of Learning Resource Systems. Their system called SMILE aims to interoperate heterogeneous information systems in the e-learning domain.

The other group [9] offers an ontology-based adaptive, personalized e-learning system, assisted by software agents on cloud storage. The proposed system integrates an e-learning application with the semantic web via domain ontology. The ontologies created for the system not only provide re-usable content for the future applications with similar purposes, but also represents the concept hierarchy structure as a domain ontology with more expressive relations.

The work [14] describes a layered architecture to manage a large amount of learning content in data grid environments, aiming at fast and precise content retrieval. The main idea of this work is to increase precision by an ontology-based semantic search, and to reduce search time by ontology-based indexing.

The [6] research suggests a framework for building an adaptive Learning Management System (LMS). The proposed architecture is based upon multi-agent systems and uses both Sharable Content Object Reference Model and semantic Web ontology for learning content storage, sequencing and adaptation. The group [2] propose a model is an automated semantic e-learning system based on BNF rules and the OWL ontology language that is capable of representing course contents using ontology.

The described information systems of storing and managing the content of education are aimed at the standard approach, such as the creating of separate objects in order to have a possibility of reusing it. At the same time the main aim of our system is to support and manage the dynamic flow of information with a possibility to restructure the system not only by the developer, but also by the user.

III. OUR APPROACH

In this section the architecture of the proposed system is explained and its components are introduced. Wiki engine is chosen for implementing the architecture since it is simple, platform neutral and threaded.

A. SCORM

The first and most important feature about this system is that we organize its structure using a network ontology model. It means that the adding of new concepts and dependences will not lead to change of structure that is very important for the teacher who is operating with the permanently updated information stream.

Secondly, to organize the teaching material in our system we use the SCORM standard as an example [1]. However, if it is necessary, it is always possible to adapt the teaching material for the use of other standard. SCORM is a set of specifications for developing, packaging and delivering high-quality education and training materials whenever and wherever they are needed [11]. In SCORM, the content packaging scheme is proposed to package the learning objects into standard teaching materials [10]. The content packaging scheme defines a teaching materials package consisting of 4 parts, that is, (1) Metadata: describe the characteristic or attribute of this learning content, (2) Organizations: describe the structure of this teaching material, (3) Resources: denote the physical file linked by each learning object within the teaching material, and (4) (Sub) Manifest: describe this teaching material as consisting of itself and other teaching material [14].

The metadata described in such standards is rather convenient for the perception of the person and can be used for the automated processing (for example, for a strict search in values of elements of metadata). However, for the execution of more composite functions, and also for the support of the independence of the expanded models of metadata we need to design the ontology model of representing the knowledge.

B. Wiki-engine

The idea of this work is to organize educational materials by an ontology approach based on the wiki engine. There many wiki-engines at the moment - for example, MediaWiki , which is developed by PHP languages, is adopted by many libraries. For our project we have chosen TWiki.

Twiki is the cross-platform open source engine, written on Perl, aimed at the creation of corporate solutions. The engine is oriented on permanently updated information stream and provides the convenient interface for its saving and updating in the system. Main characteristics of TWiki can be summarized as follows: 1) Information is divided into sections, for each section it is possible to set the access rights. 2) The full monitoring of data versions is supported. 3) Convenient and flexible reference system. 4) Simple and available web interface. 5) Easy editing through a WYSIWYG editor. 6) API for writing of extensions and plug-ins. 7) Minimum system requirements for the server and minimum system requirements for clients. 8) Embedded database. 9) Simple and fast initial setup. 10) Free license.

The main aim of TWiki is to control the flexible constantly updated information stream.

A number of studies have relied on various wiki-engines in order to construct systems that will be able to manage different types of material and documents [5]; [4]; [15]. Despite the different objectives of these systems, the overall methodology of construction them is quite similar.

The authors [13] also propose a wiki-based platform in order to rapidly generate the teaching materials. They show that a Wiki-based rapid prototyping approach to designing teaching materials for e-Learning grids is characterized by a time-saving development process, minimal human involvement, reducing redundant effort and high-quality teaching materials.

Moreover, the authors [3] have presented their research project aimed at adopting semantic web technology to draft the IEP (Individualized Education Plan) for pupils with special educational needs in school. The modelling was performed with a customized version of MoKI, the Modelling Wiki, a collaboratively mediawiki-based tool to model ontological and procedural knowledge. The customization consisted in defining ad-hoc forms to guide users in contributing to the ontology, as well as in developing specific features to browse the ontology content.

C. The design of the data structure

This paper proposes an ontological model for structuring the contents in the information system based on the SCORM standard. For structuration of data and simplification of their subsequent use, it is necessary to formalize the object basis in an ontology format and to create a group of templates for new pages. The general structure of the directory of data is designed as a meta ontology with the hierarchical type of relations and is provided in Fig. 1. Each concept of the offered meta ontology is a web template that contains formal information about itself including special metadata.

The concept of the templates is provided in Fig. 2. Each course contains the basic structured information about itself and about its components, such as subjects and general information. Each subject contains information about lectures, laboratory works, definitions, tests, etc. In turn, each of them includes texts, diagrams, pictures, links to documents, video records.

Moreover, each of these concepts is connected to the concepts of the target audience, which include the competencies of the students, that are related to the targeted group. This is shown in Fig. 2.

The web template, that is based on the ontology concept, represents a set of educational material with the description of those parts of learning objects which are included in the particular course according to the standard.By using the specified templates, the users will carry filling of the knowledge base.

TWiki exposes the minimum requirements to system: the availability of Perl and Web server. The requirements for computing capacity depend on the planned load on the system and on the number of users. The access for staff in the university will be provided via the web interface.

The main difference of TWiki from the other Wiki engines is its structural component. In the process of filling the system is not only creating the pages containing information on the courses and target audience, but is also creating the special structure in the form of the semantic ontology containing parent and child members and cross-references.

D. Scenario

TWiki provides a standard web interface for filling and editing content and background information. Each user will be able to replenish the knowledge base with the new information and new decisions.

Special templates have been created in the proposed system that are used as the basis for filling the system with learning content. These templates, as the whole system itself, are based on the application of SCORM approach - each template is a web page that contains meta information, information about the structure and the resource links. Since TWiki belongs to the category of semantic wiki engines, it contains the built-in wiki markup, as well as the ability to describe learning objects using metadata. In the process of filling the information the user enters template metadata that allows to structure and subsequently find all necessary information. Basic metadata used in our system: course, subject, resource, target audience, annotation, etc. The templates are also divided by type: lectures, laboratory work, test, general information, etc. The teacher selects a template that suits him and then puts in the data he wants to, simultaneously affixing labels and tags. Such labels are used to identify

each learning object and the target audience of students to which the educational object is intended. For example, laboratory work 1_1 (LR1_1) designed for 2nd year students of the direction of information systems (IS) and laboratory work1_2 (LR1_2) is describing the same subject, but refers to the 2nd year student enrolled in the direction of business informatics (BI). Further, while filling the page, the user creates semantic ontological links that determine the type of links between the related pages. For example, the page with Lecture 5 contains the link of type [[part of :: Mathematical Modelling]], which means that this lecture is a part of course of mathematical modeling.

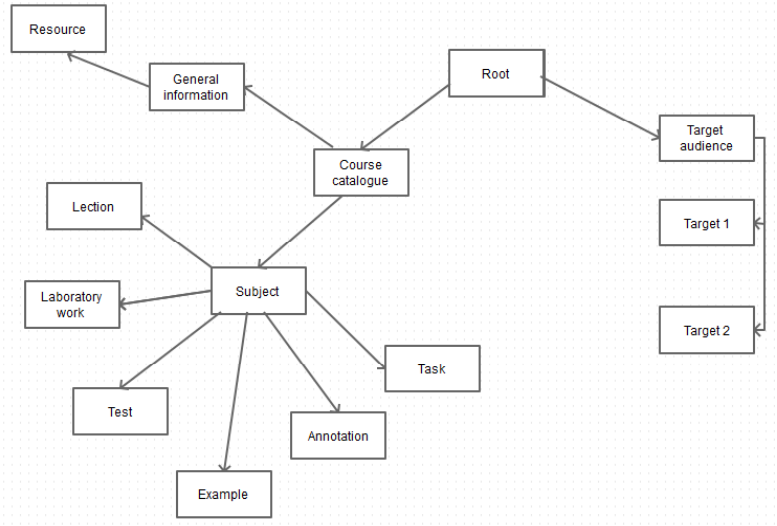

Fig. 1. Part of metaontology structure

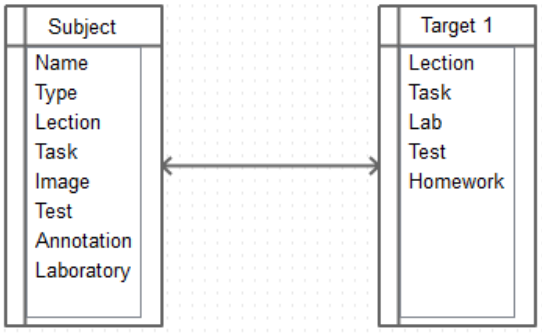

Fig. 2. Concept of the templates

When the teacher adds new documents, lectures, or even changes the structure of a course, all he needs is to simply change the desired information in the metadata and associate a new object with all others which are connected with it. This semantic Web ontology structure allows the user to make complex queries and to find necessary information not only on the coincidence of text request, but on the whole structure of the system. The teacher can search for not only child members of a certain object, but also return its parents and search for the related documents such as links, metadata, forms and information about the attached files. The ontological structure is shown in Fig. 3.

E. The knowledge base usage

In process of filling the system will save up enough data on educational process, and will be able to be used for improving and simplifying the working process of teachers in the preparation of educational material. The system search will allow the teacher to look for the necessary information about different training courses, considering the requirements of the specific target audience.

Twiki has a built-in version control system. This system allows to monitor the changes made in the system and also provides information on the users making changes, and exact time of change of documents, templates. This system

allows the user to track changes and also provides the information about the exact time when the document templates were changed.

Since one of the main objectives of our information system is not just structuring the information included in the curriculum, but also flexible updating and a selection of the available data, we have expanded the initial opportunities of TWiki by connecting an additional plug-in - the plug-in of exporting the data to an RDF ontology format.

F. The development of a data export plug-in

In the process of accumulation of knowledge of educational process it becomes possible to use it in artificial intelligence systems, such as systems of search of the optimal solution, etc. To solve this problem, it is necessary to have an opportunity to export into an acceptable format for artificial intelligence.

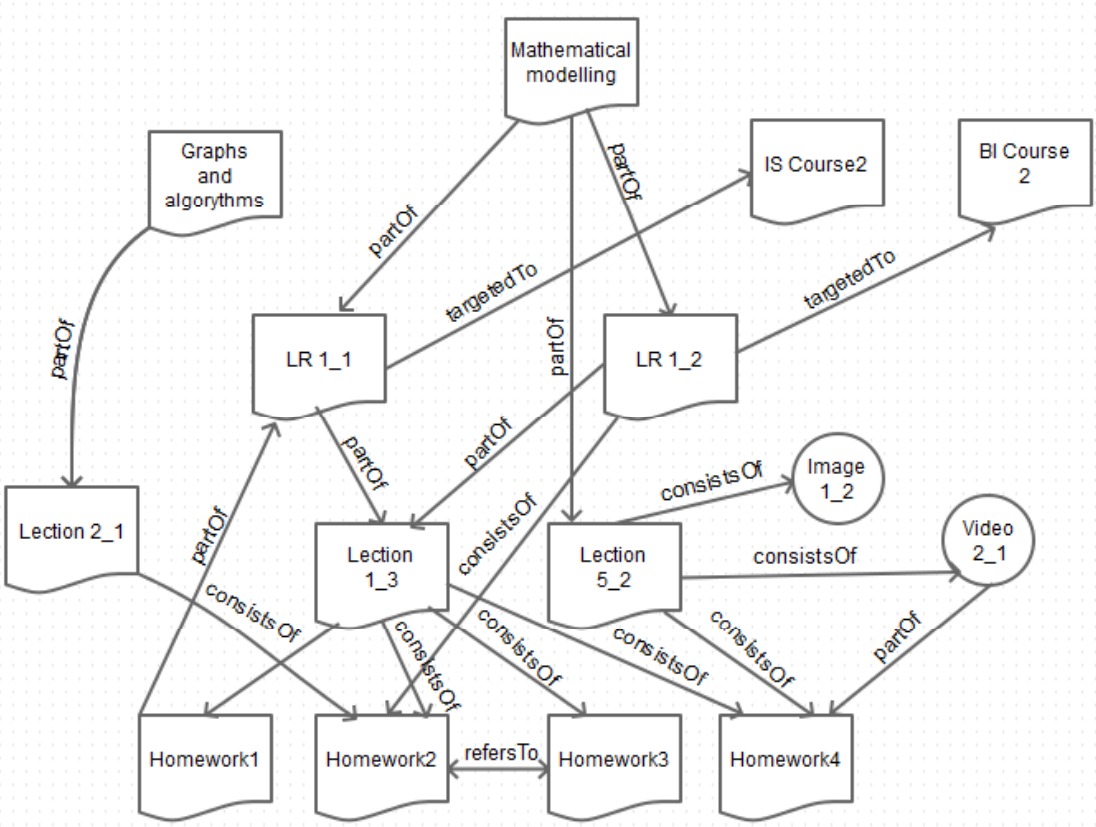

Fig. 3. Part of the implemented ontology

Respectively within our work we developed the export plug-in that allows to upload the data from the system in RDF format.

TWiki has a powerful API for creating plug-ins of any complexity and orientation. During the existence of this system, a large number of plug-ins have been created, and is stored in the open access repository. Some of these plug-ins are designed to interact with external systems.

TWiki includes built-in tools to work with third-party databases, including NOSQL databases. With the help of of the developed plug-in we request data from storage and then transfer it to an ontological format. The plugin automatically detects the type of objects and relationships between them, and uploads a formalized structure of the ontology. It is important that as the amount of data and the complexity of its structure increases, the user can add new classes and dependencies that will not affect the original structure of the system.

For submission of ontological data of our system three options of widespread formats of storage of RDF of data - XML, by NTriples and JSON were tested. Testing was performed with the use of real educational data. The speed of the system has been chosen as the main criterion.

The results of testing are given below in figures 4-7.

As we see from the diagrams, the NTriples format showed the highest performance in different requests. Respectively, we decided to use this format for data storage

on our system. However, if necessary, this plug-in gives an opportunity to upload data also in two other formats - for example, if we need the export of data in the SCORM format.

The created RDFPlugin allows to create and export ontological data, and respectively the structure of our system to the RDF format. With the help of this plug-in the additional RDF storage of our data is created. In case of making any changes, adding and deleting data in our system the information in RDF storage is updated.

IV. TESTING EXPERIMENTS

Within the development of an information system several types of testing were carried out.

The load testing of the system was held with the use of the LoadUI tools which allows to create the load of the system with the given number of the virtual users. Due to the fact that at the moment the schedulable loading doesn't exceed 30 users, the number of the virtual users was made 50.

The diagrams below show the results of the load tests. Figure 8 shows the number of the users who are simultaneously using the system and the number of requests per second. The maximum number of users was 48 and the maximum number of requests per second, slightly exceeded the limit of three.

Figure 9 presents a graph which shows that the average response time of the system under load of 48 users was less than three milliseconds. The only peak of the diagram illustrating really big time of a response in five seconds was, according to a log file, at the moment when the system turned on a garbage collector.

In general, the results of load testing gave satisfactory results. They show that, even under load, which is almost twice exceeding schedulable, the information system behaves regardless of the number of concurrent users accessing it.

V. CONCLUSION

In this paper, an ontology-based approach for building an information system for storing educational content is described. This approach gives an opportunity to effectively fill, manage and use training materials by teachers in universities. On the basis of this approach the structure of the system using the SCORM standard as a basis is designed. The system of metadata, links and summaries of templates which comprise the educational content is developed. The experimental testing results showed that the proposed system tends to cover the needs of users.

This work aims to consider the application of the presented approach in universities in order to improve the preparation of the educational material according to the needs of the teachers. Future works are oriented towards the implementation of this system at university to its test on the basis of real indices of users and further improving the system.

REFERENCES

[1] ADL SCORM, ADL SCORM (online) (cited 2008), Web: http://www.adlnet.gov/Technologies/scorm/default.aspx.

[2] A. Jehad, H. Mohammed, T. Skander, M. Mehedi, "Well-formed semantic model for co-learning", *Computers in Human Behavior*, 51 (2015), 821–828.

[3] S. Cramerottia, D. Ianes, "An ontology-based system for building Individualized Education Plans for students with Special Educational Needs", *Procedia - Social and Behavioral Sciences*, 217 (2016), 192 – 200.

[4] J. Jung, K. Kim, D. Shin, J. Park, "FlowWiki: A wiki based platform for ad hoc collaborative workflows", *Knowledge-Based Systems*, 24 (2011), 154–165.

[5] J. Jung., "Computational reputation model based on selecting consensus choices: An empirical study on semantic wiki platform", *Expert Systems with Applications, 39 (2012), 9002–9007.*

[6] M. Yaghmaie, A. Bahreininejad, "A context-aware adaptive learning system using agents", *Expert Systems with Applications* ,38 (2011), 3280–3286.

[7] N. Arch-int , S. Arch-int, "Semantic Ontology Mapping for Interoperability of Learning Resource Systems using a rule-based reasoning approach", *Expert Systems with Applications*, 40 (2013), 7428–7443.

[8] C. Ogescu, S.A. Bucharest, C. Plaisanu, D. Bistriceanu,, "Web based platform for management of heterogeneous medical data", *In Proc. IEEE Int. Conf. Autom.*, Quality, Testing, Robot.

[9] M. Rani, R. Nayak, O.P. Vyas, "An ontology-based adaptive personalized e-learning system, assisted by software agents on cloud storage", *Knowledge-Based Systems*, 90 (2015), 33–48.

[10] SCORM, SCORM best practices guide for content developers (online) (cited 2008), Web: http://www.lsal.cmu.edu/lsal/expertise/projects/developersguide/.

[11] SCORM 2004, SCORM 2004 overview (online) (cited 2008), Web: http://www.s1000d.org/events/user_forum/performance_fest2.

[12] W3C: Resource Description Framework (RDF), Web: http://www.w3.org/RDF/.

[13] W. Shih, S. Tseng, C. Yang, "Wiki-based rapid prototyping for teaching-material design in e-Learning grids", Computers & Education, 51 (2008), 1037–1057.

[14] W. Shiha, C. Yang, S. Tseng, "Ontology-based content organization and retrieval for SCORM-compliant teaching materials in data grids", *Future Generation Computer Systems*, 25 (2009), 687–694.

[15] M. Zappa, M Hoffmeistera, A. Verla, "Methodology to apply semantic wikis as lean knowledge management systems on the shop floor", Procedia CIRP, 12 (2013), 444 – 449.

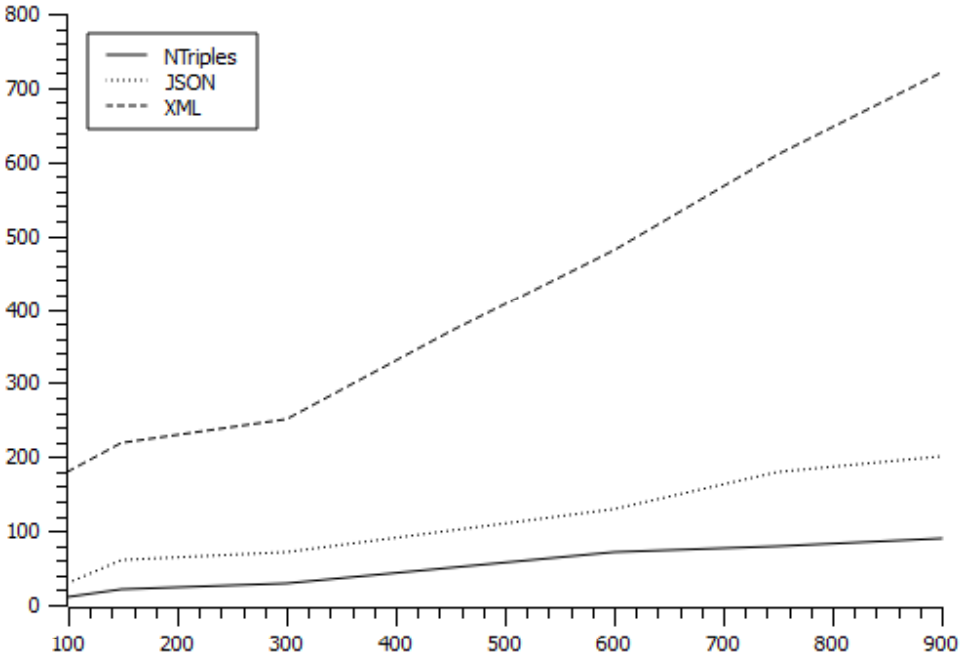

Fig. 4. File upload time

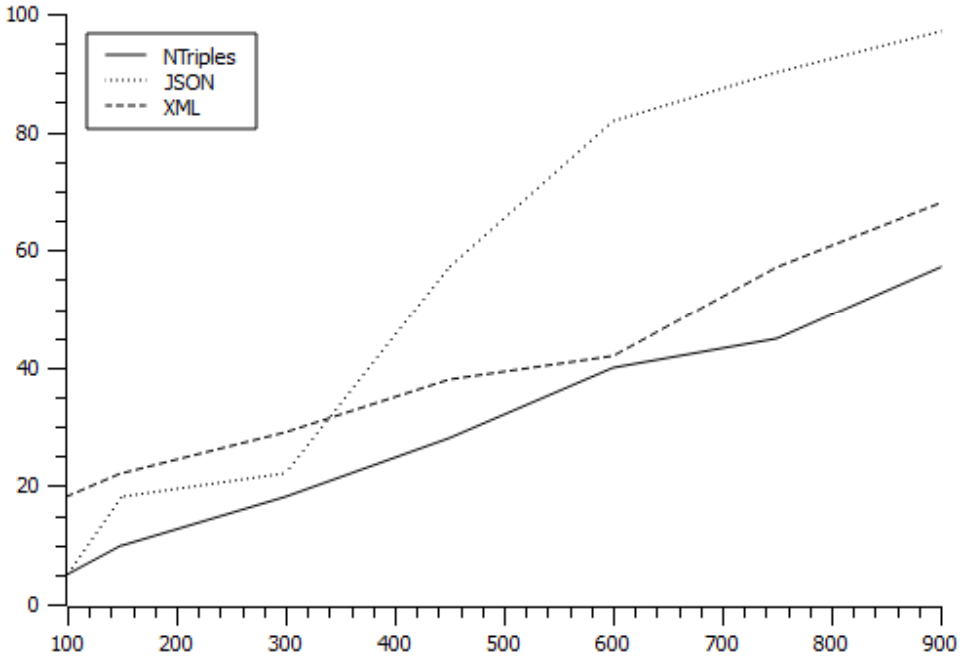

Fig. 5. Data upload time

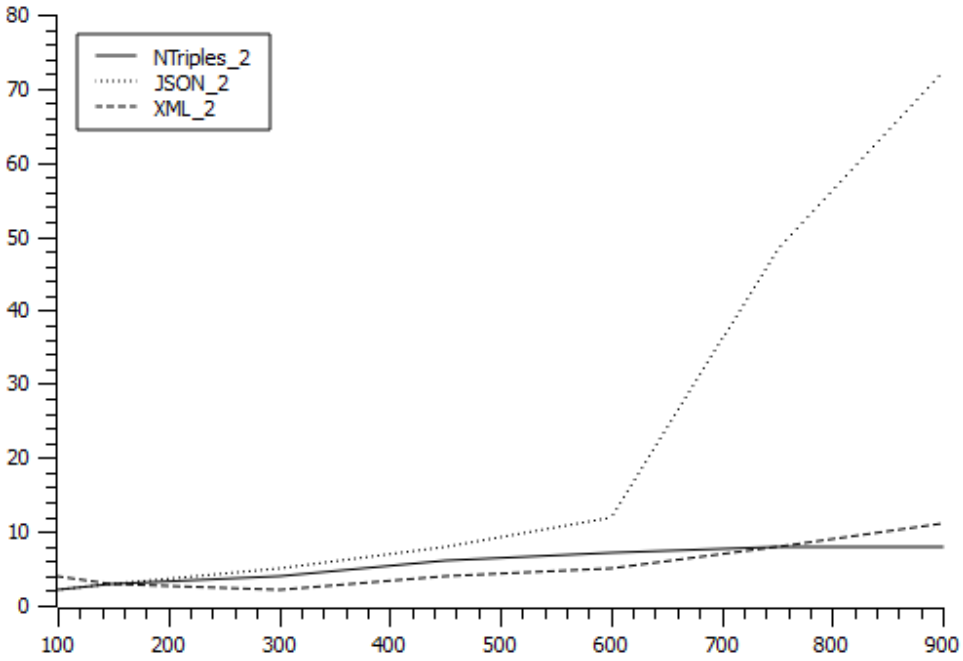

Fig. 6. Sampling time of data of one type

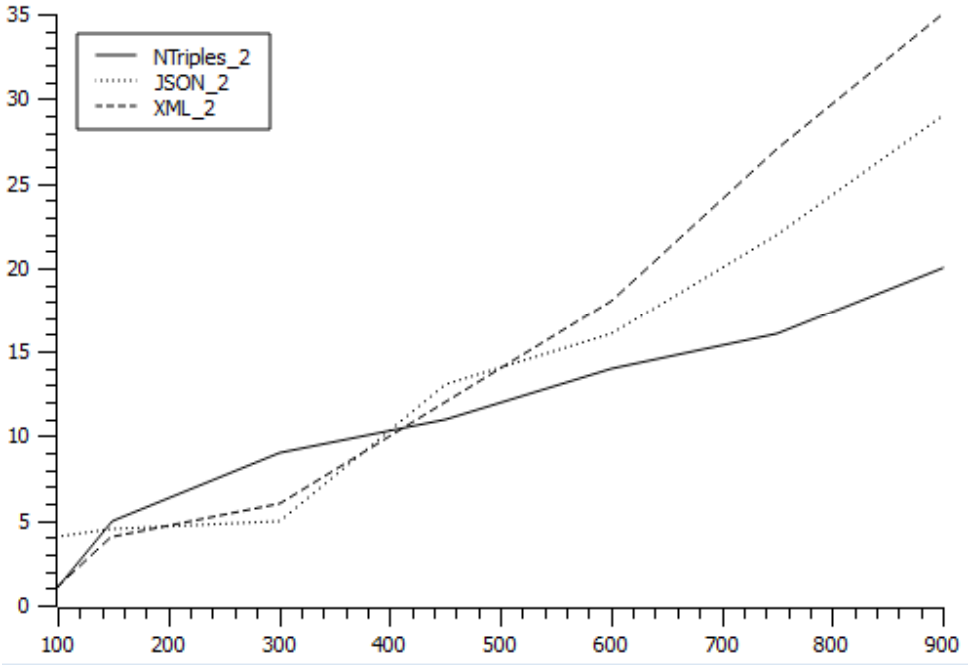

Fig. 7. Time of receiving the tree of elements

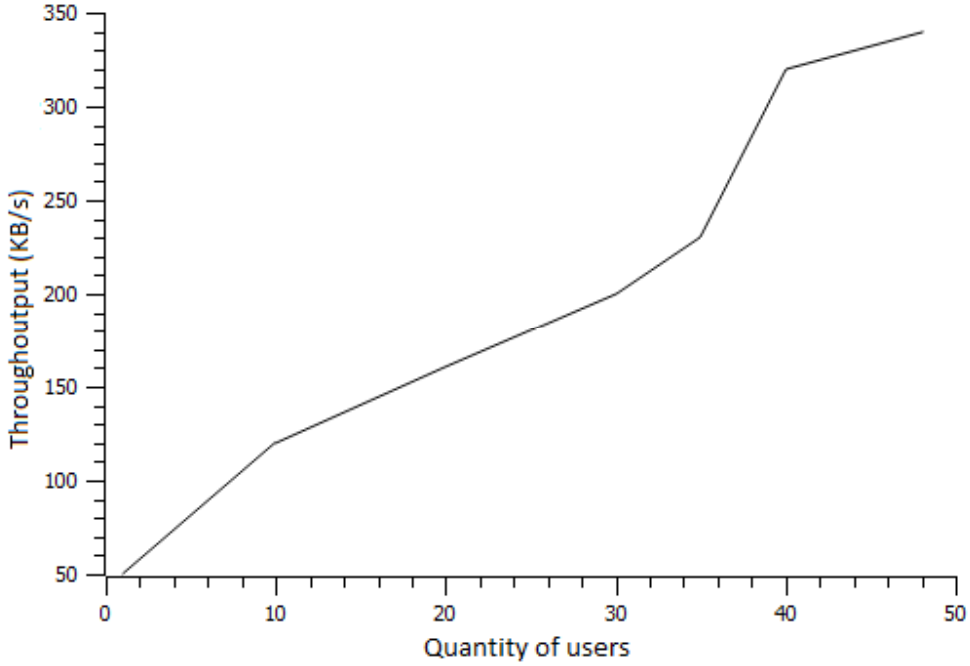

Fig. 8. Load tests diagram

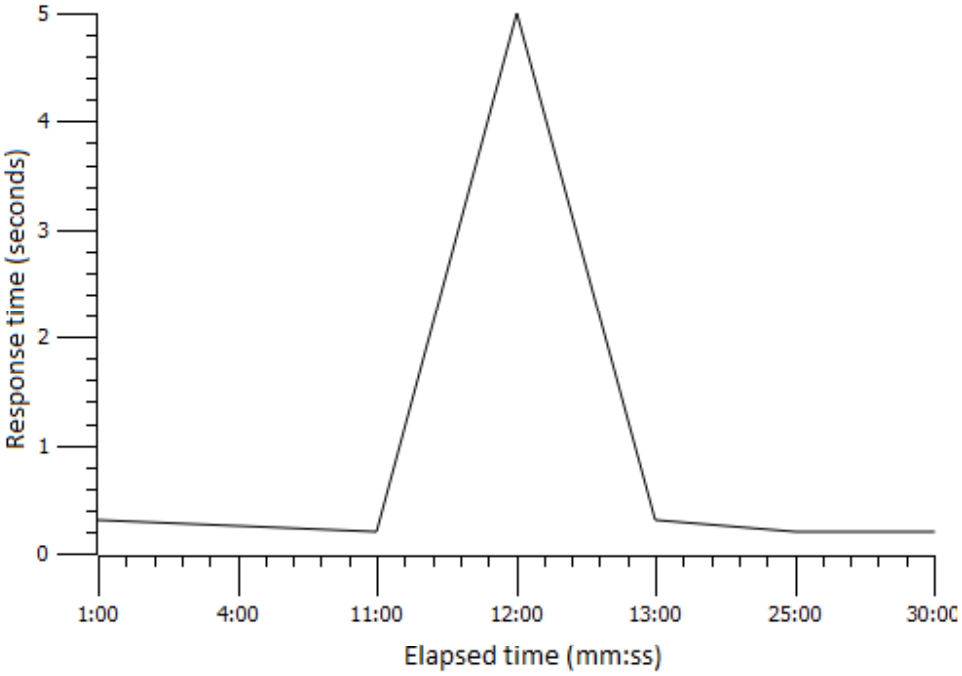

Fig. 9. Load tests diagram