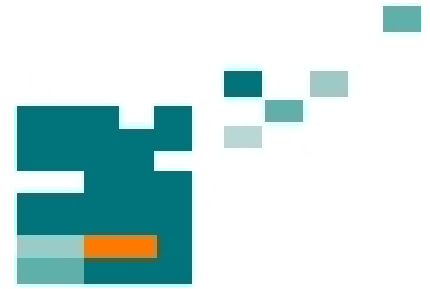


54. IWK
Internationales Wissenschaftliches Kolloquium
International Scientific Colloquium



**Information Technology and Electrical
Engineering - Devices and Systems, Materials
and Technologies for the Future**



Faculty of Electrical Engineering and
Information Technology

Startseite / Index:

<http://www.db-thueringen.de/servlets/DocumentServlet?id=14089>

Impressum

Herausgeber: Der Rektor der Technischen Universität Ilmenau
Univ.-Prof. Dr. rer. nat. habil. Dr. h. c. Prof. h. c.
Peter Scharff

Redaktion: Referat Marketing
Andrea Schneider

Fakultät für Elektrotechnik und Informationstechnik
Univ.-Prof. Dr.-Ing. Frank Berger

Redaktionsschluss: 17. August 2009

Technische Realisierung (USB-Flash-Ausgabe):
Institut für Medientechnik an der TU Ilmenau
Dipl.-Ing. Christian Weigel
Dipl.-Ing. Helge Drumm

Technische Realisierung (Online-Ausgabe):
Universitätsbibliothek Ilmenau
[ilmedia](#)
Postfach 10 05 65
98684 Ilmenau

Verlag:  Verlag ISLE, Betriebsstätte des ISLE e.V.
Werner-von-Siemens-Str. 16
98693 Ilmenau

© Technische Universität Ilmenau (Thür.) 2009

Diese Publikationen und alle in ihr enthaltenen Beiträge und Abbildungen sind urheberrechtlich geschützt.

ISBN (USB-Flash-Ausgabe): 978-3-938843-45-1
ISBN (Druckausgabe der Kurzfassungen): 978-3-938843-44-4

Startseite / Index:
<http://www.db-thueringen.de/servlets/DocumentServlet?id=14089>

VIRTUAL INFORMATION ENVIRONMENT IN SCIENCE OF LAW BASED ON ONTOLOGICAL KNOWLEDGE MODEL

*Vasyl' Tatsyi¹, Anatolij Getman¹, Oleksandr Sokolov², Mikola Shvets³, Sergii Prylypko¹,
Stanislav Ivanov¹, Volodymyr Karasiuk¹, Oleksandr Lugoviy¹*

¹Yaroslav the Wise National Law Academy of Ukraine, Kharkiv, Ukraine

²National Aerospace University, Kharkiv, Ukraine

³Legal Informatics Scientific-Research Centre of the Academy of Legal Sciences of Ukraine, Kyiv, Ukraine

ABSTRACT

Modern lawyers work with large volume of dissimilar text information. But accessibility of this information, search organization of necessary law information, information support for all kinds of law activity in Ukraine are still imperfect. In the Yaroslav the Wise National Law Academy of Ukraine Information Technologies Center (Kharkiv, Ukraine, the leading high educational establishment in training qualified lawyers) join with NAU "KAU" information department (Kharkiv, Ukraine, the leading high educational establishment in training specialist in the sphere of informational technology) start work on creation of perspective law informational systems' elements. The technical base of informational system prototype development is the National Law Academy's multi-service informational network. Principles of artificial intelligence, particularly ontological approach to law information organization are the program base of informational system organization. The training process of lawyers in the Academy will be the sphere of using this system.

Index Terms – law information, local network, informational services, knowledge base, ontology, information search, web-interface, training system.

1. INTRODUCTION

Law system in Ukraine is working extremely ineffective. There are many reasons for such situation: its becoming in the conditions of democratic state; intensive life of normative base, which produces its instability; incapability of main corps of practical lawyers to use modern informational systems effectively; lack of appropriate information in accessible electronic format. Except above-mentioned subjective factors there is an objective ones: lawyers' work connected with working and using of great mass of unstructured and dissimilar text information, which is hardly formalized. For example, effective legislation of Ukraine (Presidential Acts, acts and rules of

Verhovna Rada, directions of Cabinet of Ministers and others) consists of about 300 hundred documents. At the same moment, law activity, its practical, scientific and educational forms based on its means, among them law documents are essential. From the cybernetic point of view, any law subsystem is complex informational system which has initial information about the law task on the entry and information about adopted documents in the form of law documents on the exit of law subsystem [1]. Information value, which is necessarily used in the process of effective law activity so various and large that even experienced lawyers meet with difficulties taking their decisions and waste a large amount of time.

For solving mentioned problems and giving the opportunity for prospective professional lawmakers to get modern experience in using law informative resources, in the Yaroslav the Wise National Law Academy of Ukraine, the local computer network has been built on the base of optic -fiber lines [2] and join with the National Aerospace University "KAU" subject-oriented databases are being formed for supplying effective practical work of lawyers, supporting educational process and scientific work. Scientific researches direct at the development of principles of law information organization, creation of educational system, approbation of elaborated approaches in the Law Academy's teaching process and practical using in the legal clinic, which works with the Law Academy participation. At the same time, main efforts are concentrated on designing the binding technologies of various kinds of law information, automatic procedures of filling base of knowledges, creation of web-interface system.

2. TECHNICAL ENVIRONMENT WORK WITH INFORMATION IN THE YAROSLAV THE WISE NATIONAL LAW ACADEMY OF UKRAINE

In the Yaroslav the Wise National Law Academy of Ukraine (Kharkiv) during last 3 years the multi-service informational network (LAN) was built.

System approach gave an opportunity to inculcate the up-to-date technologies of information storage and retrieval into the nets and to provide present-day informational services for all the users of the network. Just therefore the network is named as multi-service. , Information Technologies Center is responsible for all the organizational questions that are concerned with creation of new network segments, informational servicing of users, information security in the network and others. The conception of the automation of educational and scientific work in the academy requires that every student, student with master's degree, post-graduate student, research worker or lecturer must have access to the information resources of the network from any work place. Work places of the network are placed at the chairs, dean's offices, reading rooms, in the sports complex, in specialized computer rooms and in the rooms of student's hostels.

The overalls of the network are the following: the whole length is more than 40 kilometers; more than 800 workstations, server HP ProLiant DL380 with two processors Intel Xeon 3GHz is used. From the overall number of workstations more than 400 are computers in the rooms of student's hostels. The capacity of the disk array is 540 Gb with the ability to widen it up to 1Tb. Also for the optic information transformation there are 4 D-Link mediaconverters, 2 guided switchboards, 20 mediaconverters in different buildings and 18 usual switchboards. The created environment is able to provide all the existent up-to-date and perspective information services. The infrastructure of the LAN is shown at the figure 1. At the scheme separate objects of the academy and their connections as they are located in the town are shown. The main cables' capacity is from 48 to 4 fibers. Some of the hostels are located more than 15 km from the main building.

3. INFORMATIONAL SERVICES IN THE MULTI-SERVICE INFORMATIONAL NETWORK.

Requirements that are formulated regarding learning process' organization at present-day level will be given gradually within the bounds of learning portal of the academy. As the experience of the e-learning development has showed e-portals are at the heart of its organization in Ukraine. Main tasks during the portal's development are accumulation and systematization of information resources with the defined thematic direction and the complex of services' organization that will guarantee efficient use of the information resources. Although the portal is in its developmental stage some of the servers are already available for students, lecturers and deans within the network, namely:

- automation of the students' module examination;
- records management of the students' progress and attendance;

- provided access to the library's website and catalogue from every work place;
- access to the FTP-server of the educational resources. The educational resources are organized in accordance with disciplines that are taught by all the chairs;
- supplying with the IP- telephony services;
- review of programs on digital TV (including programs from the academy television centre);
- e-mail receiving and sending;
- telephony supporting by means of local area network (LAN) from the workstations in the academies' sub-units and students' hostels;
- using services of files exchange (ICQ) by the users within the LAN;
- providing with the Internet access;
- access to the InfoBase of legal documents that is updated twice a day from the server of Verkhovna Rada of Ukraine;
- review of the images of some web cameras that are placed in the academy;
- conditions' preparation for realization of elements of the distance learning in the academy;
- the other auxiliary and supplied services.

4. ONTOLOGICAL APPROACH FOR CREATION KNOWLEDGE BASE (KB) OF LEGAL INFORMATION

The research task is a creation of theoretical foundations of the information storage and retrieval system on basis of present-day linguistic technologies of maintenance the learning process in the legal education. Creation of information storage and retrieval system means creation of KB as the basis of information storage and retrieval. In this case ontology is the most convenient form of the KB. There are different treatments of the "ontology" term that depends on context and purposes of its using. In the paper we suggest the definitions given in [3, p.7] "Ontology is the specific fact that is created for the meanings expression that are interpreted in the shared vocabulary"

According to the formal representation most of sources propose the wording that is similar to [4, p.285], where ontology is presented with the triple:

$$O = \langle X, R, \Phi \rangle$$

where X is finite aggregate of concepts;

R is finite aggregate of the ratios between X elements;

Φ is finite aggregate of interpretation functions that are specified on the concept and/or ratios of the ontology O.

The ontology complexity depends on every aggregate's content. In [5] 5 levels are marked out:

- glossary;
- taxonomy;

- thesaurus (taxonomy with terms);
- conceptual system with any number of ratios;

- completely computer-based system.

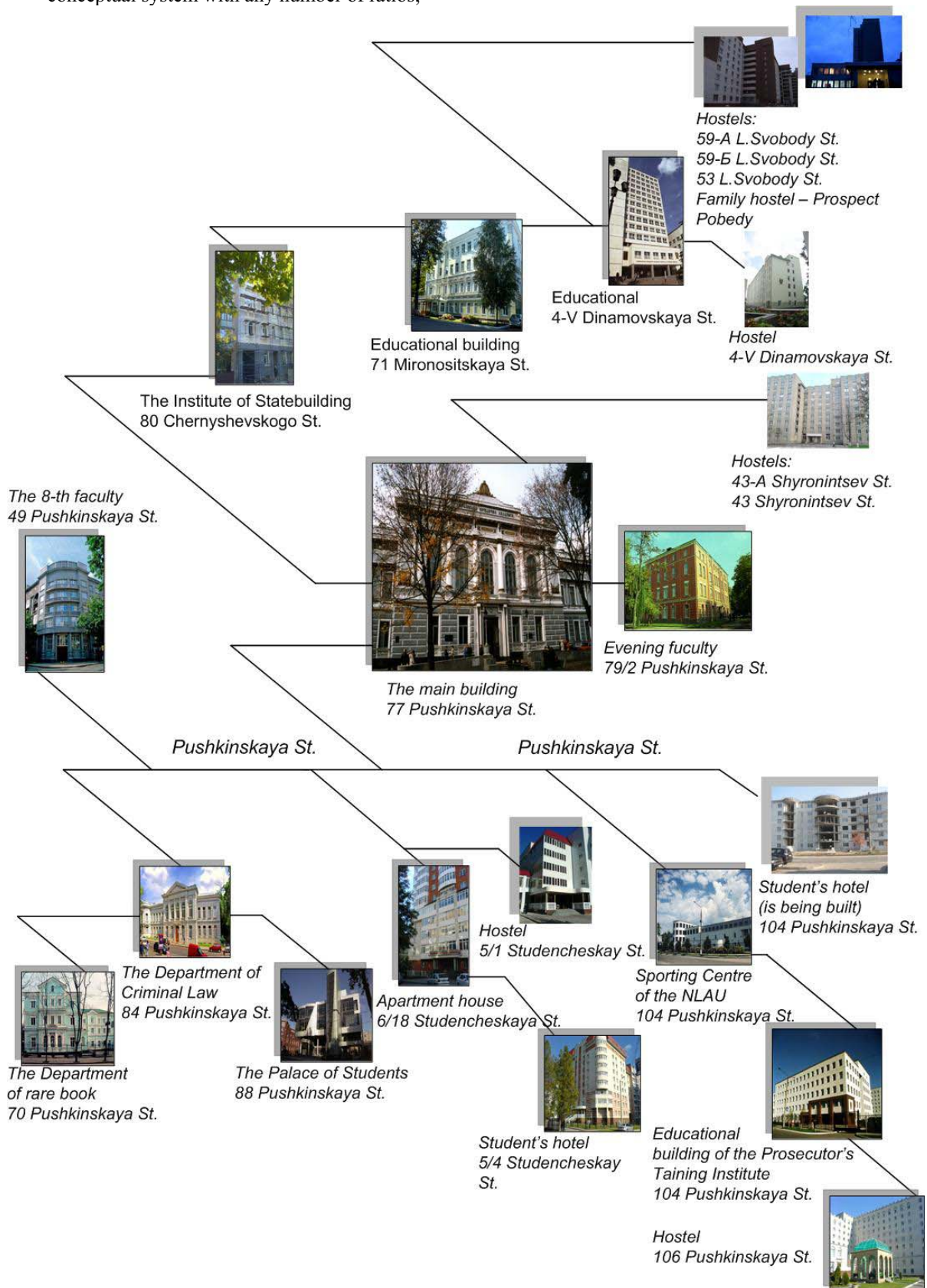


Figure 1. The infrastructure of multi-service network of Yaroslav the Wise National Law Academy of Ukraine

For solution of the problem the fourth variant was selected in with subset Φ is empty and X and R are

arbitrary filled. Reason of rejection from more complicated ontology's presentation level is necessity

to fill the ontology from texts. Also we were obliged to refuse a system that is based on hierarchy of classes with inheritance of the property sets and mechanism of samples (like in Protege) as to mark out the denoted aspect in the data domain turned out to be problematically even in manual mode.

On the whole the developed system is presented as a system of associated definitions with the ability to enhance not only quantity of used concepts but also quantity of used bunches.

5. PROGRAMMING PACKAGE

Taking into consideration the fact of multiple and various usage of the programming package we decided to divide it into 4 program products fulfilling the following goals:

- filling ontology database and indexation of texts;
- reviewing the indexed texts with the further transition based upon the connection of the notions mentioned, and searching for a text using a key word/phrase;
- interactive cooperation with a person with a teaching and knowledge checking goals;
- arranging the text for the further inserting it into a database.

The goal of filling the base with data is realized by a narrow circle of experts. That is why for intensifying the process of writing and extending the programming opportunities the graphic part is presented as Java Application. The parts for looking through and teaching are presented as Web Applications for simplifying the access to the system.

5.1. The Structure of Database for Saving Ontology

The offered structure includes: the data of possible variants of spelling, connection between the concepts, texts concerning the concepts and connections among them, mentioning the sources of the texts and their structure. The structure of the database is presented in Figure 2.

Various forms of spelling of terminology and connections of the concepts in the texts are saved in eight tables. All the words used for spelling the terminology and relations of the concepts are put into the table "words". The words united into the groups of synonyms are kept in the table "swords". Each of the concepts (the table "concs") and relations (the table "rel") are presented in the word combinations (the table "ssents") which in their turn are presented as itemized packages of groups of synonyms.

The relations among concepts are kept in the tables "rrel" and "cons_rrel". The first table keeps the information about types of the relations and expiry

dates; the second one keeps the information about the interconnected concepts.

The tables "text_conc" and "text_rel" contain the information about concepts, their relations and sentences. When working with concepts one should remember about the advantage of keeping the concept location in the sentence presented in numbering the words belonging to spelling the concepts. The texts are presented in the set of sentences. The sentences are kept in the table "texts" which also keeps the number of the source and section which this sentence belongs to, and the location of the sentence in the section. The numbers of sections, their hierarchy, number in the parent section and the number of the source are kept in the table "sects". The attributes of sources of texts are kept in the table "tstc" which is connected with the table of authors of sources.

Thus, though the presented model of saving information of concepts and their relations based upon simple word-by-word regime of saving might in some cases slow down working with database but it provides the considerable level of flexibility when saving the information or searching for the data.

5.2. Working Scripts

At this point one of the four parts of programming package, the appendix for filling the database with ontology knowledge and indexation of texts was completed. The preformatted texts saved in xml format, word concepts, relations and specific relations among groups of concepts are presented as input data. The user can enter them through screen forms.

For filling each of the type of data there is a corresponding form.

For applying for a word concept or the relations there are forms "Concept Editor" and "Editing the Word Combinations, Synonyms and Words" which provide the access to the names and the structure of the elements of the word concept.

For setting the relations among the word concepts there is a form "Editing the Relations among the Concepts". This form provides information about all the relations of the chosen element and simplifies search for establishing the connection between the main and subordinate groups. This form also helps add and edit the relations between the earlier-created concepts.

For indexing the texts there are two forms: "Indexation of Texts" and "Editing the Sources". The first form allows providing the full circle of processing of the downloaded text, from marking the fragments concerning the concepts to setting new concepts and relations among the group of concepts. The second form allows setting data of sources of texts and downloading the texts into database from preformatted xml files.

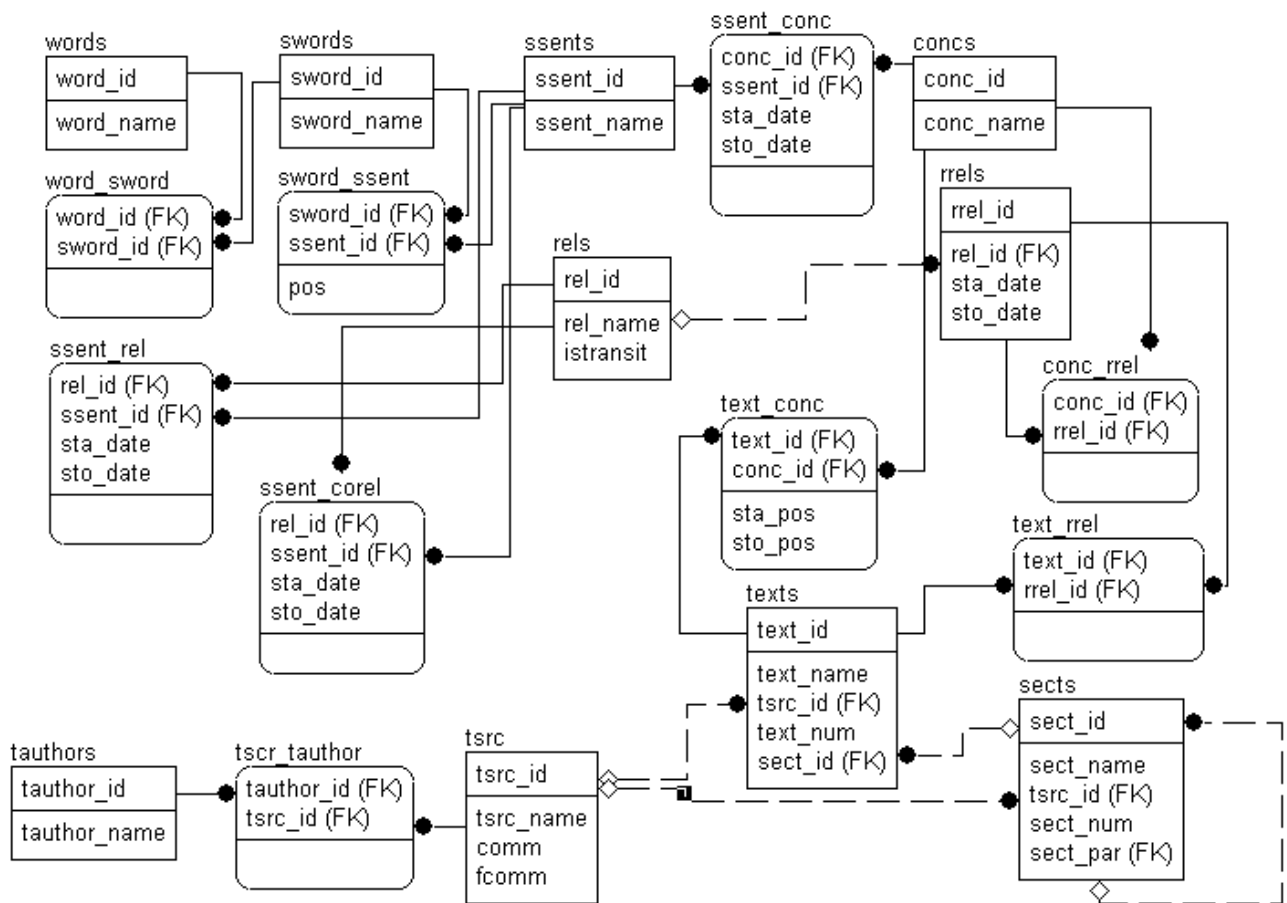


Figure 2. ER-diagram of Ontology Database

5.3. Programming Classes Applications for Filling Database

For making the programming process more comfortable and for further improving the program product we used the division of products into groups of classes by the character of their usage. The mentioned division is presented in Figure 3.

The list of the main classes of application with explanation of their purposes follows further.

Package core:

DatabaseConnector – this class provides an access to database by the means of JDBC. The name of the file containing the configuration of JDBC is transmitted to the constructor. The methods for carrying-out SQL query are realized in this class and also trivial queries: end, commit, rollback and others with the aim of reducing constants in the code which have been compiled.

GlobalVariables – class contains global variables used in the application.

Language – class contains constants for receiving lineal meanings of messages in different languages.

Parameters - class contains constants for receiving external form of application and parameters of lead-in to database from the file of tuning.

Start – basic (main) class of application.

Package data model:

Idstring – line with ID element.

IdstringWDate – line with ID and date of start and finish.

SentElemData – class of performing word belonging to the notion (graphic presentation of italicized notions and (links in the sentences).

Package data model.construct of:

ConcWordLineStruct - class of performing one of the branches of the tree in the verbal form of notion or link.

ConcWordTreeStruct – class for recording tree of performing in the verbal form of notion or link.

SsentTreeStruct - class for recording tree of performing in the verbal form of notion or copula (link).

SwordTreeStruct - class for recording tree of performing in the verbal form of notion or link.

Package data model.guidata:

SentPanelData – class for keeping words and operations corresponding to it for SentPanel.

Package datamodel.textstruct:

Sentence - class for recording sentences of the text into the general structure of performing the text.

TextSection - class for recording sections of the text into the general structure of performing the text.

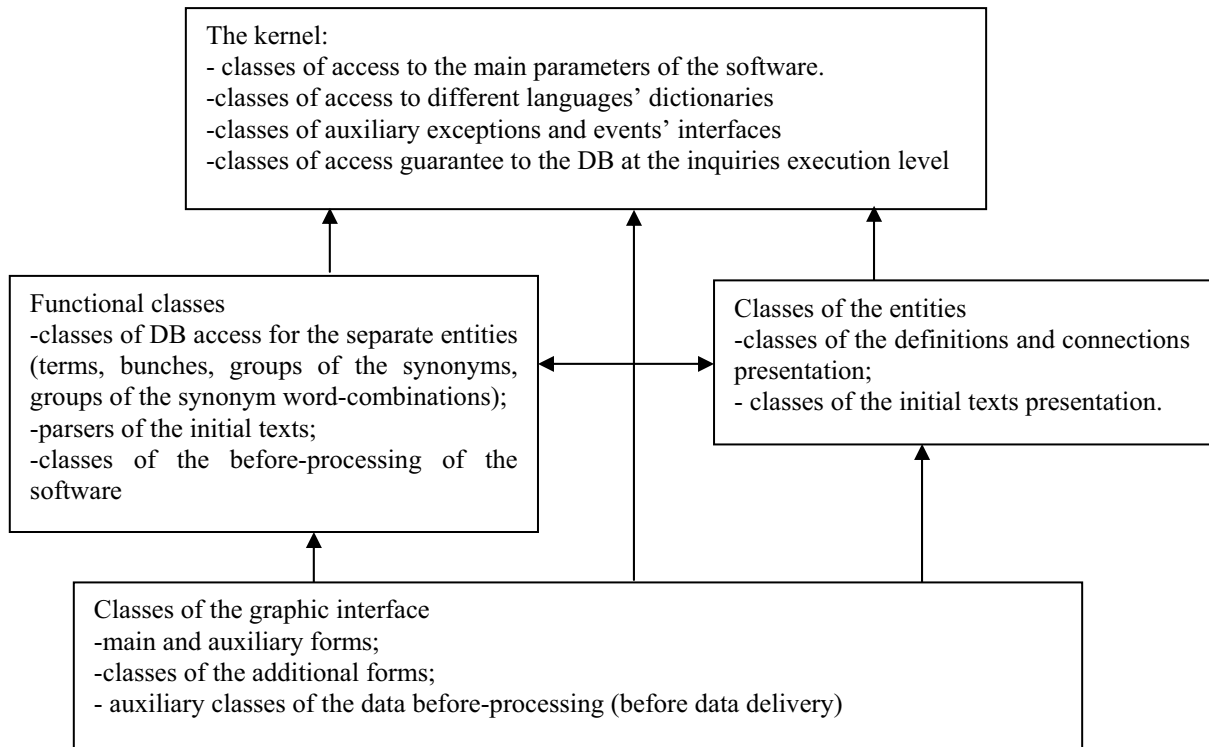


Figure 3. Application of classes dependences.

TextSrc - class for keeping database about the source of the text.

TextStr - class for keeping text and its structure.

Package functionality:

SpellDict - class for .hunspell, to provide access to linguistic dictionaries.

Package functionality.database:

DBConcEdit - organization of access to database for receiving and recording verbal performing of notions and links.

DBLang - class providing keeping current language of text's analysis and performing notions and links; access to the names of the languages in the base.

DBRrelEdit - organization of access to database for receiving and recording links between groups of notions.

DBText - organization of access to database for receiving and recording texts (full structure performed by TextStr).

DBTextSrc - organization of access to database for receiving and recording data about the sources of the texts (authors, type of commentary).

Package functionality.database:

TextParser - class realizing analysis of xml file with initial text into recording the content in class TextStr, and also writing down to the base.

Package functionality.database:

CheckWord - form for choosing element (word-combination, group of synonyms, word) with reviewing (checking) its structure.

ConcWin - form for creating and editing verbal performing of notions and links (independently connects with base).

CorrectWord - form of choosing "vocabulary" meaning of the word (all kinds of vocabulary forms and correcting forms).

EntityWin - form of task and editing verbal performing synonymic word-combinations and group of synonyms.

InputWord - auxiliary (subsidiary) form of input of the line.

LangAction - class of providing change of language of performing texts in all set forms.

MainWindow - form of basic (main) window of program.

RrelWin - form of creating and editing links between groups of notions.

SentPanel - widget, providing representation of the sentence in the form of set of panels with word and corresponding it actions.

TextWin - form for indexation of texts with the possibility of adding of new notions of links.

TSrcWin - form for task and editing data about sources of text (independently connects with base) (editing of the text is not provided).

WordPanel - widget, providing representation of word and corresponding it actions.

5.4. Graphic interface of user

Application provides light change of language of the interface at the expense of keeping all user's messages in separate file.

As mentioned above application contains 5 basic (main) forms presented on the pictures4-8.

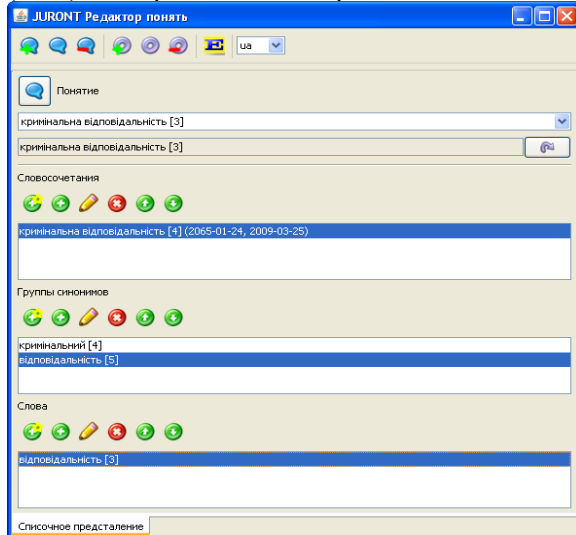


Figure 4. Form of input verbal performing notions and links.

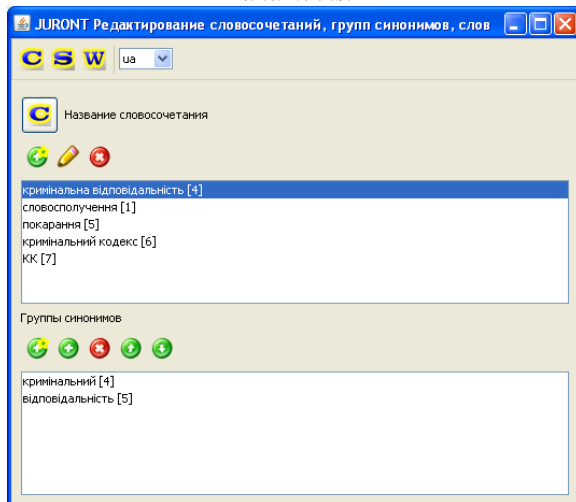


Figure 5. Form of input verbal performing word-combinations and groups of synonyms.

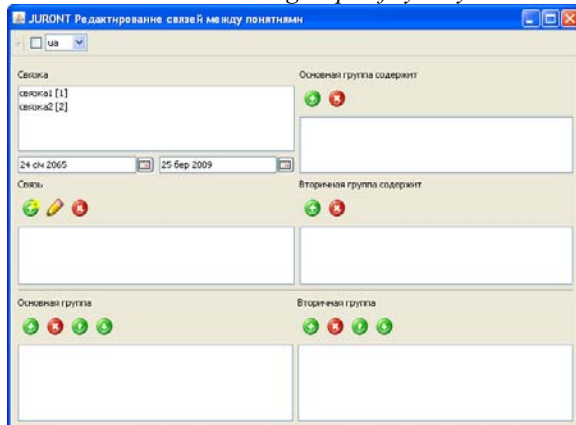


Figure 6. Form of tasks of links between groups of notions.

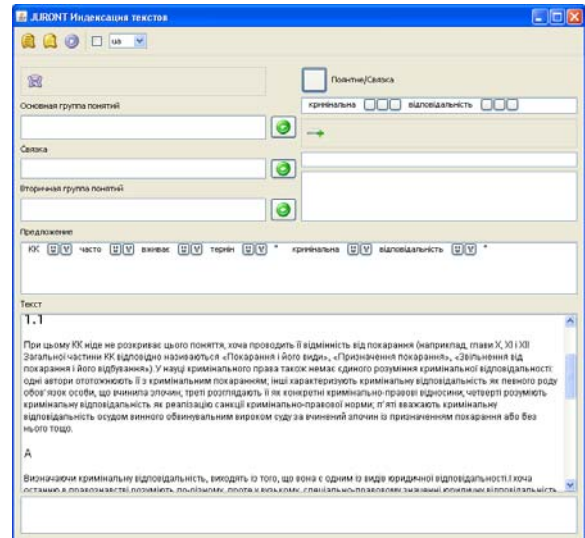


Figure 7. Form of processing texts.

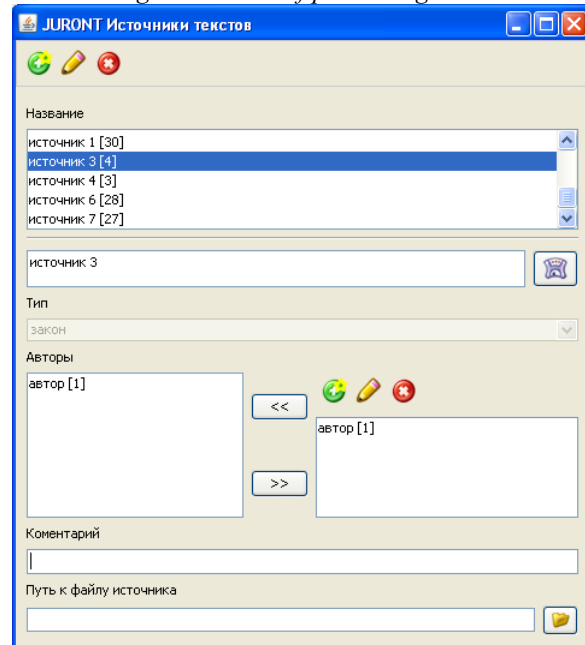


Figure 8. Form of loading text to the base and tasks of characteristics of the source of the text.

6. CONCLUSIONS

Existing informational and program systems do not provide lawyers by all kinds of work with legal information.

Working out of knowledge-oriented informational system for educational purposes is carried out at Yaroslav the Wise National Law Academy (Ukraine, Kharkiv). The ontological approach is on the base of this system. According to this approach database of legal information is represented by the notional system with arbitrary composition of relations. The program complex of this system consists of four program products which provide all conditions (regimes) of this system. Web-interface of user and automatized regime of work with base of knowledge

and automatized filling of ontology from accessible textual documents are provided.

Full realization of this project will allow all users in multi-service informational network of Law Academy to have access to the resources of informational and educational system of new generation.

7. REFERENCES

[1] Gorevoy L.E., Shvets M.Ya. Computer system of informational and analytic provision of law-making and law-enforcing activity. Kyiv, Ukraine. Parliamentary Edition (in Ukrainian) 1998.

[2] Getman A., Ivanov S., Karasiuk V.. Informative providing of modern education // Scientific Information for Society – from today to the Future. Abstracts. 21st International CODATA Conference. October 5-8, 2008. Kyiv, Ukraine. National Technical University of Ukraine “Kyiv Polytechnic Institute”, p. 89 – 90.

[3] Konstantinova N.S. Ontologies as systems of keeping knowledge / N.S.Konstantinova, O.A. Mitrophanova // All –Russian contestual selection of reviewing articles on prior direction "Informational telecommunicational systems", 2008. – 54 p.

[4] Gavrilova T.A. Bases of knowledge of intellectual systems / T.A. Gavrilova, V.F. Khoroshevskiy – St.P.: Piter, 2000. – 383 p.

[5] Dobrov B.V. Linguistic ontology on natural sciences and technologies for supplement in sphere of informational search / Dobrov B.V, N.V. Lukashevich – [Electronic resource]: http://fccl.ksu.ru/issue_spec/docs/oent-kgu.doc
[13.01.09]