Creation of Terminological Resources for Record Management, Terminology Science and Information Retrieval as Well as Terminological Dictionary

SERGEY V. SILKOV, DR.
Republic of Belarus - Minsk - 220101, G.Yakubov's str., 28-380
e-mail: SILKOV_S_V@MAIL.RU

Creation of Terminological Resources for Record Management, Terminology Science and Information Retrieval as Well as Terminological Dictionary

ABSTRACT

This article touches on various approaches to terminological information resources (TIR) which are created by this author. TIR «InfoLio&Rec» is intended for researches in area of scientific disciplines of archives and document sciences, records management, spaces of Universum of Keeping (archives, libraries, museums), knowledge management, information law, e-government. This TIR is described in details as one of the information research system developed on the basis of normative and legal acts operating in UNO, ISO/IEC, EU, Belarus, Russia, Ukraine and other countries regulating deeds and record flows. TIR «Terminus», being a computer aided system for studies in terminology of terminology science and information retrieval, is also dealt with. Basic principles, traits and structure of the dictionary (also created by this author) «Archivology and documentology in Belarus, Russia and Ukraine. Contemporary terminology» are revealed.

Creazione di risorse terminologiche per la gestione documentale, scienza terminologica e recupero delle informazioni, nonché dizionari terminologici

SINTESI

Questo articolo tocca diversi approcci alle risorse informative terminologiche (TIR) che vengono creati dall'autore. TIR «InfoLio & Rec» è destinato a ricerche nel settore delle discipline scientifiche degli archivi e delle scienze dei documenti, gestione dei record, spazi di Universum di Keeping (archivi, biblioteche, musei), gestione della conoscenza, diritto dell'informatica, e-government. Il TIR è descritto in dettaglio come uno dei sistema di ricerca d'informazione sviluppato sulla base di atti normativi e giuridici che operano in UNO, ISO/IEC, Unione europea, Bielorussia, Russia, Ucraina e altri paesi che regolamentano atti e flussi di record. TIR «Terminus», essendo un sistema computerizzato per lo studio della terminologia della scienza terminologia tratta anche del recupero delle informazioni. Vengono descritti i principi fondamentali, i tratti e la struttura del dizionario (creato anche da questo autore) «Archivology e documentology in Bielorussia, Russia e Ucraina. Terminologia contemporanea».

Ustvarjanje terminoloških virov za pisarniško poslovanje, terminološko vedo in priklic podatkov kakor tudi terminološki slovar

IZVLEČEK

Prispevek se nanaša na različne pristope k terminološkim informacijskim virom (TIV), ki jih je ustvaril avtor. TIV "InfoLio&Rec" je namenjen raziskavam na področju arhivskih in dokumentacijskih znanosti, pisarniškega poslovanja, ohranjanja dediščine (arhivi, knjižnice, muzeji), upravljanja z znanjem, informacijskega prava, e-uprave. Natančno je opisan TIV kot eden izmed informacijsko raziskovalnih sistemov, ki je bil razvit na osnovi normativnih in pravnih aktov, ki se nanašajo na upravljanje z dokumentarnim in arhivskim gradivom, in ki jih uporabljajo v OZN, ISO/IEC, EU, Belorusiji, Rusiji, Ukrajini in drugih državah, ki nadzorujejo pretok dokumentov. Prispevek obravnava tudi računalniški sistem "Terminus", ki pomaga pri študiju terminologije v terminološki znanosti in pri priklicu podatkov. Podaja osnovna načela, značilnosti in strukturo slovarja "Arhivologija in dokumentologija v Belorusiji, Rusiji in Ukrajini. Sodobna terminologija", ki ga je razvil avtor.

Создание терминологических ресурсов и терминологического словаря по управлению документацией, терминоведению и поиску информации

РЭЗЮМЭ

В настоящей статье затронуты различные подходы к терминологическим информационным ресурсам (ТИР), которые создаются автором статьи. ТИР «InfoLio&Rec» предназначен для исследований в области научных дисциплин архивологии и документологии, управления документацией, пространств Мира Хранения (архивы, библиотеки, музеи), управления знаниями, информационного права, электронного правительства. Этот ТИР подробно описан как информационная исследовательская система, разработанная на основе нормативных правовых актов, которые приняты в ООН, ISO/IEC, ЕС, России, Беларуси, Украине и других странах, и регулируют действия с документами и их движение. Также рассматривается ТИР «Terminus» - информационная система для проведения исследований в терминологии терминоведения и поиска информации,. Выявлены основные принципы, особенности и структура создаваемого автором словаря «Архивология и документология в Беларуси, России и Украине. Современная терминология».

As known, research on the development, classification, codification and standardization of documentary, communicational etc. terminology is of significant scientific importance. The transformations in these fields of science directly impact corresponding systems of terminology, hence «systematical research of various professional lexicons» (Hejchuk... [et al.] 2004) is an important task.

Terminology of each science or activity reflects a wealth of its accumulated professional knowledge (Nazin... [et al.] 1969). The introduction of information technology throughout modern society speeds up the development of records management standards and terminology unification by the international organizations. Adaptation of numerous foreign terms into national terminology puts its consistency to the test: leniency towards illogicality of adopted terms might results in new Tower of Babel (Duchein 1965), but neglecting the terminology is no less dangerous. Research on the records management terminology allows examining in detail the theoretical and historical aspects of its integrity and consistency issues, its development and evolution, build-up of systems of terminology, adoption, issues of terminography (as the theory, methodology and practice of creation of terminological reference works such as terminological dictionaries and databases), etc.

Current extralinguistical situation dictates the development of terminological information resources (TIR) which:

1. would become established means of documenting the terminology, knowledge accumulation and professional discourse,
2. would provide complete and accurate information to support the selection and correct usage of the terms (including general and applicability information, legal status, references to definitions in legislation/regulations, etc.),
3. could become a scientific basis for the standardization of systems of terminology.

My TIR InfoLio&Rec I was created as reference and retrieval system on the topic «The Government and the Law». The TIR contains 8900 terms and definitions derived from the review of the state-of-the-art of science and IT; from the laws, regulations, standards and guidances etc. published by the Republic of Belarus, other countries and international organizations; and from other sources on museum, library and archival science, records and information management, information law, e-government and related areas.

Translating the texts related to several areas of knowledge is difficult due to the lack of contemporary dictionaries reflecting the recent changes in the terminology (Hejchuk... [et al.] 2004). At present InfoLio&Rec contains the Russian, Belarus and English terminology, while Ukrainian terminology is being added.

Relevance and usefulness of the InfoLio&Rec are due to:

1. variations in the terminology frameworks of the specified areas of knowledge;
2. need to streamline and standardize; the lack of comprehensive terminological dictionaries for these areas; and absence of an information system capable of providing full information concerning specific systems of terminology;

ATLANTI • 22 • 2012 • n. 2

Sergey V. SILKOV: Creation of Terminological Resources for Record Management, Terminology Science and Infor-mation Retrieval as Well as Terminological Dictionary, 157-161

3. different interpretations of some terms by scientists and practitioners;
4. need to harmonize national systems of terminology with international ones, to fasten new terms and to purge obsolete terms from the systems of terminology.

Main goal of «InfoLio&Rec» is *continuous supporting of the processes of standardization of professional terminology* in order to increase efficiency and stability of social communications and management, i.e. providing methodological support for the development of consistent and unambiguous terminology and corresponding TIR. It also supports retrieval of term's definition and any other necessary reference information on terminology-related objects and processes.

InfoLio&Rec can be used as a basis for the development/improvement of specialized databases or directories. However, it is not intended as a replacement of all the existing specialized databases or directories.

Achieving InfoLio&Rec's main goals requires:

1. Identification of all the numerous terms in use in the specified subject domains;
2. Separation certain subject domain's terms from the terms of related domains;
3. Identification of factors influencing and impacting the system of terminology;
4. Identification of a subset of terms which is necessary and sufficient for subject domain's progress;
5. Development of hierarchies of concepts;
6. Development of clear definitions of the terms;
7. Validating reasonableness of keeping polysemy and synonymy;
8. Elimination of violations of the general and special language rules.

Technical rules and regulations recording approved practices and decisions (e.g. codes, standards, guidances, etc.) are the important form of accumulation of knowledge. They are used as a primary means of ensuring interoperability of information systems, especially in Internet-based communities. Therefore International Standards on records management developed by ISO and IEC, as well as laws, regulations and guidances published by the UN, the European Union, the Republic of Belarus, Russia and other countries, were used as primary sources. It is worth noting that no single source can provide comprehensive coverage of all the subject areas covered by the TIR. Besides, the majority of means of fixing terminosystems in the specified areas are incomplete and inconsistent.

Nowadays the specialists should be able to select specific information system supporting the transition from «simple» terminological resources (e.g. computer-stored manuals, phonetic and training systems, dictionaries and glossaries) to advanced processing and retrieval systems, machine translation systems etc. (Belyaeva 2008). Every TIR should support searching of terms by keyword, topic, entity, document; or identification of a term on the basis of detailed information about it. If InfoLio&Rec's user fails to find the required terminology, (s)he can be re-directed to Internet resources and/or key publications in the area of terminology and corresponding areas of knowledge.

Interaction of a professional lexicon with general language system is reflected in the processes of terminologization, re-terminologization, de-terminologization (Hejchuk... [et al.] 2004). Therefore TIR should be introduced into educational processes. Difficulties in memorizing the definitions will be compensated in this case by the availability of clear images (models) of definitions and their refreshing in memory. The system of terminological training in the subject areas supported by InfoLio&Rec includes the terminology in a number of languages, the list of abbreviations and special notations. The suggested technology of terminology training is based on identification of a group of terms corresponding to each dictionary term. Such a group generates the thesaurus/micro-dictionary supporting hierarchy of terms, relations between them and (or) certain operations. The role of special dictionaries «as means of recording and representation of knowledge is rising now. (...) Thesaurus-type dictionaries quite possibly could become a core of artificial intellect systems. (...) The dictionaries can also be used as a means of reconstruction and research of the process of development of human knowledge.» (Grinev-Grinevich 2009).

Another my TIR - «Terminus» - covers the terminology of terminological science and information search. «Terminus» pays special attention to the terminology associated with basic concepts such as: *term, terminology, system of terminology, terminosphere*, etc.

The base of sources, dissertations and dictionaries was formed on the basis of the keywords presented by following root morphemes (for the Russian language): *граммат*, дефини*, знак*, индекс*, ключ*, когит*, когнит*, лекс*, лингв*, логик*, логич*, определен*, поиск*, прагма*, семант*, семио*, синта*, систем*, слов*, словар*, словн*, текст*, термин*, фраз*, энциклопед*, язык**

As V.I. Glumov (from prof. B. Golovin's scientific school) suggested, when subject domains are represented using minimum keyword list, it is necessary to specify synonyms. Quality of these lists is essential; otherwise the search results will be incomplete and not accurate enough. This is demonstrated by processing of the concepts *compound term, external term structure, underlying term structure, semantic multiplier, search accuracy, search completeness*, etc.

The sources were classified as follows:

1. Monographs and papers published by classical and modern linguistic schools (von Humboldt's, Austrian, French-Quebec, Geneva's, Danish, Prague, Polish, etc.), Soviet and national Belarus schools, the CIS countries (more than 4600 references);
2. Dissertations published by Soviet and national Belarus schools, the CIS countries (more than 8370 references);
3. Glossaries, lexicons, terminological collections, dictionaries, directories, encyclopedias (more than 8200 references).

The references to the works of key scientists and researches will be included into separate lists sorted alphabetically by researcher's name. Since TIR covers a number of subject areas, the necessary references to the sources on philosophy of a science, consciousness, mathematics, physics, psychology, jurisprudence also are included.

List of recommended e-resources includes:

1. The legal dictionary of terms and definitions at the National law portal of Belarus;
2. Topical glossary of Russian language;
3. Explanatory dictionaries authored by Dahl, Ushakov and Ozhegov;
4. Wikipedia;
5. Oxford's and Webster's explanatory dictionaries;
6. Help system of the Xrefer network;
7. The World of Dictionaries» web-portal (including business, biography, encyclopedic, historical, political, sociological, philosophical, financial, economic, law dictionaries, dictionaries of names, computer slang, logic, symbols and synonyms);
8. Additional e-resources.

By analogy with TIR InfoLio&Rec the section on terminology training is included.

TIR Terminus works with the information converted in formats .DOC, .HTM, .PDF is possible. TIR InfoLio&Rec works only with the information converted in a format.HTM.

The InfoLio&Rec and Terminus are registered in the Republic Belarus State Register of Information Resources (The Certificate on inclusion of the encyclopaedic electronic web-dictionary of modern terminology «InfoLio&Rec» in the State register of information resources of Belarus № 1100900613 from 12.01.2009; The Certificate on inclusion of the encyclopaedic electronic web-dictionary of modern terminology «Terminus» in the State register of information resources of Belarus № 1161001057 from 01.10.2010).

The basic principles of dictionary which I designing with intermediate name as «Archivology and documentology in Belarus, Russia and Ukraine. Contemporary terminology»:

1. The dictionary should keep continuity with *historically developed traditions of archival and document sciences* in Belarus, Russia and Ukraine, operating documents: regulations, positions, instructions, etc.;
2. There should be no rupture from fundamental achievements of classical document sciences
3. Dictionary should consider developed market realities of the Belarus, Russia, Ukraine: accents in the dictionary are essentially displaced towards electronic archival and document sciences.

ATLANTI • 22 • 2012 • n. 2

Sergey V. SILKOV: Creation of Terminological Resources for Record Management, Terminology Science and Infor-mation Retrieval as Well as Terminological Dictionary, 157-161

4. Dictionary is uneffective in a separation from national lawa, standards on systems of electronic archives and records flow, interaction protocols, etc.

Roughly, the dictionary should, conditionally speaking, to be subdivided into sections:

1. basic concepts (*documentary information; information system; information access; documenting; information management*, etc.);
2. concept of separate kinds of documents and their carriers (*original; copy; version; extract; report; file;* etc.);
3. terms and compound of terms, which define the organizations, the structural divisions which are carrying out information and documentary activity, which basic function are storage and processing of documents (*archives; current reports; records centre;* etc.);
4. concepts which opens of record management processes and their tools (*appraisal; criteria of value of reports; lists of documents with periods of storage; reports inventory; control dictionary; classification; thesaurus; indexing; storage*, etc.).

Sources, literature

Ludmila BELYAEVA, *The automated workplace of the philologist: principles of the organisation and componental structure,* In *Actual problems of applied linguistics. Proceedings. Part 2. ed. A. Zubov*, Minsk 2008, p. 27-28 (in russian)

Michel DUCHEIN, *A propos de la terminologie archivistique,* In *La Gazette des Archives, 1965. N. 50, p. 158*

Sergey GRINEV-GRINEVICH, *Introduction in terminography*, Moscow 2009, p. 153-154 (in russian)

Roman HEJCHUK... [et al.], *Bases of creation of terminological dictionaries*, Bialystok 2004, p. 7 (in russian)

Ivan NAZIN... [et al.], *Soviet archival terminology*, "Soviet archives", 1969, n. 1, p. 31 (in russian)

SUMMARY

The TIR InfoLio&Rec was created as reference and retrieval system on the topic «The Government and the Law». The TIR contains 8900 terms and definitions derived from the review of the state-of-the-art of science; from the laws, regulations, standards and guidances etc. published by the Belarus, other countries and international organizations; and from other sources on museum, library and archival science, records and information management, information law, e-government and related areas. At present InfoLio&Rec contains the Russian, Belarus and English terminology, while Ukrainian terminology is being added. TIR Terminus covers the terminology of terminological science and information search. «Terminus» pays special attention to the terminology associated with basic concepts such as: *term, terminology, system of terminology, terminosphere*, etc. Roughly, the dictionary should, conditionally speaking, to be subdivided into sections: 1) basic concepts (*documentary information; information system; information access; documenting; information management*, etc.); 2) concept of separate kinds of documents and their carriers (*original; copy; version; extract; report; file*; etc.); 3) terms and compound of terms, which define the organizations, the structural divisions which are carrying out information and documentary activity, which basic function are storage and processing of documents (*archives; current reports; records centre*; etc.); 4) concepts which opens of record management processes and their tools (*appraisal; criteria of value of reports; lists of documents with periods of storage; reports inventory; control dictionary; classification; thesaurus; indexing; storage*, etc.).