Paper

# Personalized Knowledge Mining in Large Text Sets

Cezary Chudzian, Janusz Granat, Edward Klimasara, Jarosław Sobieszek, and Andrzej P. Wierzbicki

*National Institute of Telecommunications, Warsaw, Poland*

**Abstract—The paper starts with a discussion of the concept of knowledge engineering, in particular ontological engineering. Consequently, the paper presents assumptions accepted as a basis for a group research on a radically personalized system of ontological knowledge mining, relying on the perspective of human centered computing and combining ontological concepts of the user with an ontology resulting from an automatic classification of a given set of textual data. The paper presents a pilot system PrOnto that supports research work in two aspects: searching for information interesting for a user according to her/his personalized ontological profile, and supporting research cooperation in a group of users (Virtual Research Community) according, e.g., to a comparison of such personalized ontological profiles. The paper concludes with suggestions concerning diverse applications of ontological engineering tools and future work.**

***Keywords—human centered computing, knowledge engineering, ontological engineering, personalized ontology.***

## 1. Introduction

During last decade, a special importance in telecommunications and Internet services achieved *data mining* or *knowledge mining* in large data sets describing such services; related terms are called *knowledge management*, *knowledge engineering* or even *knowledge science*. However, *knowledge science* touches philosophy, and *knowledge management*, even if of computer science origin, is today treated as a part of management science; therefore, we shall rather use the term *knowledge engineering* in its broad sense, extending it beyond its classical academic sense of artificial intelligence and learning algorithms.

A research group in National Institute of Telecommunications concentrates on knowledge engineering for over ten years, together with basic research on such disciplines as mathematical logic, multiple criteria decision theory, diverse optimization and statistical methods, also ontological engineering; all these theoretical aspects serve, however, as the basis of development of tools of knowledge engineering, in particular knowledge mining in large sets of data.

Applications of these tools relate to diverse problems. They might consist in diverse data and knowledge mining services for telecom operators, or using advanced statistical methods to analyze diverse indicators of the development of informational society in Poland or in Mazovia region. However, this paper concentrates on applications of ontological engineering to support of knowledge mining, research and knowledge management.

We must add still one explanation. Classical methods of ontological engineering concentrate, similarly as typical work on artificial intelligence, on an automation of knowledge mining from large textual data sets, while the preferences of the user might be taken into account, but typically in a limited extent. The character of the work presented in this paper is different and results from our practical experience in applying data and knowledge mining. We assume a *sovereign position of the user* – explained more specifically in further text – and concentrate on a *radical personalization* of ontological user profile that might use, but should not be dominated by the results of automatic analysis of large sets of textual data[1].

## 2. Knowledge Engineering and Tacit Knowledge

Experience in applying knowledge engineering tools shows that knowledge mining is aimed not only at finding *logical relations between data*, but as well at discovering *tacit knowledge hidden in large sets of data* and correlated with *tacit knowledge of the user*. We apply here the concept of *tacit knowledge* on purpose, although it denotes usually[2] *preverbal* (difficult to express in words) knowledge hidden in human mind, see [6]–[10].

However, preverbal knowledge is contained also in large data sets, even in textual data sets, and the goal of knowledge engineering is to discover such knowledge – not only

[1]The paper describes results of work in a project called in Polish *Projekt Badawczy Zamawiany "Usługi i sieci teleinformatyczne następnej generacji – aspekty techniczne, aplikacyjne i rynkowe", grupa tematyczna i: Systemy wspomagania decyzji regulacyjnych: Wykrywanie wiedzy w dużych zbiorach danych telekomunikacyjnych (Requested Research Project "Next Generation Services and Networks – technical, application and market aspects", Theme Group i: Systems Supporting Regulatory Decisions: Knowledge Mining in Large Telecommunication Data Sets)* and is a modified version of longer Polish texts [1], [2].

[2]Usually but not exclusively, since there is also tacit knowledge in the *intuitive intellectual heritage of humanity* including *synthetic a priori judgments* [3] and *hermeneutical horizons* (see, e.g., [4]) expressing essential intuitive beliefs propagated by educational systems, as well as *emotional heritage of humanity* including between others *collective unconsciousness* [5]) together with it's parts – myths, archetypes, etc., but also all artworks, say, the emotional load of all films. Hence tacit knowledge can be contained not only in the mind of a single human being, see [6].

in an algorithmic and automatic way, but also with the cooperation or even under the guidance of a human user.
In a broad understanding of knowledge engineering we can distinguish several parts of it:

I. Narrowly understood artificial intelligence and automatic learning engineering.

II. Discovering knowledge (including tacit knowledge) in large data sets, data and knowledge mining.

III. Text processing engineering, including ontological engineering, but also textual knowledge mining.

Part I is described by many books, see, e.g., [11]. Part II relies partly on the tools developed in Part I, but uses also much broader diversity of tools: statistical, decision analytical, etc., and includes to a larger extent the requirements and participation of human users. Part III aims usually at finding or selection of textual explicit knowledge and uses tools of *ontological engineering* and *semantic Web* as well as network *search engines*; in applications, however, decisive is an interpretation of the selected textual knowledge by a human user, hence according to the user's tacit knowledge or hermeneutical horizon [4], [12], [13].
*Ontological engineering* is also related to *knowledge management*, see, e.g., [14]. The term *ontology* was borrowed from philosophy, where it means *theory of being* (see, e.g., [15]); computer science interprets differently this term as a classification of entities and words representing them. In information technology, we treat today the term *ontology* as an enriched taxonomy, vocabulary with a hierarchy and other (logical, semantic) relations of terms. A significant development of ontological engineering occurred during last two decades, related to the concept of *Semantic Web* and based on the assumption that contemporary WWW network contains (or will soon contain) knowledge corresponding to all intellectual heritage of humanity, thus advanced information technology tools should be able to extract essential part of this knowledge in form of an universal ontology[3].
Ontologies play today, when treated as tools of representation and shared understanding of knowledge about diverse domains, important roles in many applications, such as development of information systems, organizing the content of Internet pages, categorizing commercial products, standardizing vocabularies in given fields, see, e.g., [16]–[19]. However, there are diverse controversies also in ontological engineering, related to several opposite approaches to the construction, application and interpretation of ontologies. There are many methods of constructing ontologies, see, e.g., [20]; we can speak about constructing *lightweight ontologies* with a simple hierarchical structure, or *heavyweight ontologies* including more detailed logical and semantic relations between terms. We can also speak about constructing *local ontologies* characterizing terms used by a local group of researchers or even by a research discipline or a cultural sphere (the same term, such as *ontology*, might have different meaning for different disciplines), as opposed to *universal ontologies* trying to represent all knowledge contained, say, in WWW network. We can also construct an ontology *from scratch*, through *reuse*, or *automated* (using automatic methods of ontological engineering), see, e.g., [19]; the last distinction is not quite precise, since good ontological engineering tools are always *semi-automated*, assume some interaction with the human user that constructs ontology with their help, while an essential problem is the extent and character of this interaction, discussed in detail below.
As the most advanced in ontological engineering, the works of Standard Upper Ontology Working Group (SUO WG) are often cited, aimed at "*forming an upper ontology whose domain is all of human consensus reality*" together with related CYC ontology (see, e.g., [21]. This is an interesting attempt to build a universal vocabulary, but many doubts can be voiced, e.g., to the use of the term "upper" (who is upper – human or network and computer?), or to local applications of such vocabulary (local meaning might not correspond to the popular meaning in the Internet).
Another subdivision of the methods of ontology construction relates precisely to the role of a human constructor of the ontology. If we assume that it is human constructor who should be sovereign and "on top", then we should speak about *top down* way of ontology construction as starting with experience and intuition (as well as emotions) of a human expert or a group of them, while *bottom up* way of ontology construction should denote an automatic construction based on broad textual content. Thus, the "upper" ontology of SUO WG is actually a universal bottom up ontology that might be difficult to apply locally, because it does not take into account the tacit knowledge of a local group of experts.
This distinction is related also to a technical and evolutionary theory of intuition [22], [23] that uses the contemporary knowledge of telecommunications and computer science to show that the use of language (and logic) by humans simplified at least ten thousand times[4] perception and reasoning that was originally immanent (using all senses). This resulted in a tremendous *surplus of brain* that is used in diverse ways, in tacit knowing and tacit knowledge, in intuitive reasoning, existential and transcendental thinking. If only less than 0.1% of neurons in our brains is needed for logical thinking and verbal argumentation, than human intuition can be much stronger (even if still fallible) than

[3]This assumption is debatable, see footnote 2 above and [6] on the role of tacit knowledge in intellectual heritage of humanity, as well as further discussion of the reasons of radical personalization of individual ontological profiles.

[4]The broadband needed for transmission of vision is at least 100 times larger than the broadband needed for transmission of voice, and the computational complexity of processing such large data sets is nonlinear; assuming quadratic increase of complexity gives results close to a lower bound. Therefore, *a picture is worth at least ten thousand words*. Thus, when we developed speech in the evolutionary development of humans, we made a tremendous evolutionary shortcut and obtained *a surplus of brain* (some philosophers call it *surplus of mind*): only less than 0.1% of our brain cells is needed for verbal communication and rational reasoning.

logical argumentation. This, however, implies the need of a *radical personalization* of ontological profiles of users of ontological tools, relying on an increase of the role of personal intuition when defining such profiles; such radical personalization is consistent also with the trend to *human centered computing*.
Such is the perspective that motivated us to search for new approaches to ontology construction (from scratch or by reuse, with lightweight structure, combining top-down and bottom-up, semi-automated approaches) for a local group of researchers. Originally, in the *Theme Group i: Systems Supporting Regulatory Decisions: Knowledge Mining in Large Telecommunication Data Sets* of the Requested Research Project we planned a broader application of such ontological approach to support regulatory decisions on telecommunication markets, but a cut of funding forced us to limit the application to a local research group in telecommunications, affiliated at the National Institute of Telecommunications.

# 3. Results of the Work on Knowledge Mining

### *3.1. Preliminary Investigations*

Initial investigation involved cooperation with IIASA (International Institute for Applied Systems Analysis) and JAIST (Japan Advanced Institute for Science and Technology, School of Knowledge Science), see, e.g., [24], [25]. A broad survey of literature has shown that there are papers suggesting a combination of bottom-up and top-down methods of ontology construction [26] but not specifying how to combine them. In [24], [25] we proposed the use of *hermeneutic reflection* (expert reflection on the structure of local ontology), of *organizational reflection* (expert reflection on the organizational structure of a research institution); we also considered the use of *mind mapping* to stimulate the intuitive top-down construction of upper layers of an ontology by the user; the lower layers might result from a bottom-up approach and ontology matching tools might be used to combine them.
We also compared diverse available tools of ontological engineering and developed a Polish language modification of the system OntoGen. OntoGen (http://ontogen.ijs.si/) is an open source tool for semi-automated bottom-up text mining and ontology construction. We tested this system on publications of our National Institute of Telecommunication with satisfactory results, see [1], [2], [27], [28]. However, the main result of this preliminary work was an idea how to construct a radically personalized user's ontological profile, leading to the concept of PrOnto system.

### *3.2. Radically Personalized User's Ontological Profile*

We started with an analysis of an important dichotomy in search of textual information in the network. There are two opposite classes of such search problems (and some mixed problems in between):

– searching for an answer to a well defined question of the user (*information retrieval*);

– searching for information interesting for the user, but rather loosely defined (*information filtering*).

Traditional search engines combined these functions to some extent, today we observe a trend to separate them. More important for supporting research is the second class that requires, however, a specification of user's preferences. Such specification can be implicit, resulting from an analysis of the history of behavior of the user (which is a popular tool of supporting internet commerce, with a long own history – see, e.g., [29], or explicit, in the form of a set of keywords, key phrases, or even a simple ontology (which again can be constructed from scratch by the user, or be influenced by the history of user's behavior). Both implicit and explicit specification of user's preferences can be modified for supporting research (see, e.g., serwis CiteULike), but explicit specification makes it possible to preserve the sovereignty of the user while constructing a radically personalized user's ontological profile.
Such a profile (which might be called also a *perspective*, or a *horizon* of the user) is assumed to consist of three layers.

- An upper layer of *concepts* $c \in \mathbf{C}$, defined by the user and treated as her/his intuitive entities (they might be later interpreted logically, but with utmost caution, because, e.g., the concept *Markov chains* can actually mean *these aspects that are now interesting for me in the theory of Markov chains*).

- A lowest layer of *keywords* or *key phrases* $k \in \mathbf{K}$, either defined by the user or by a bottom-up ontological tools (they will be later the main connection of the radically personalized profile with classical ontological tools).

- A middle layer of *relations between* $\mathbf{K}$ and $\mathbf{C}$, or *importance coefficients* $f \in \mathbf{F}$ of a keyword for a concept, defined by the user (later they might be also modified by the history of user's behavior, but the user should be sovereign in their specification), interpreted either as weighting coefficients, or subjective probabilities, or fuzzy logic membership values, or aspiration levels for multiple criteria ranking of documents with respect to the ontological profile, see below.

The radical personalization consists in assuming that only the lower layer $\mathbf{K}$ is responsible for collaboration with bottom-up ontological engineering tools. The middle layer $\mathbf{F}$ and the upper layer $\mathbf{C}$ might form together with the lower layer a kind of personalized ontology (used, e.g., to support cooperation in a research group), but the user is sovereign in using her/his intuition when modifying these two higher layers.

# 4. Prototype System PrOnto

### 4.1. A general Structure of PrOnto

Generally, PrOnto system supports research work of a group of users (Virtual Research Community, VRC) using a radically personalized user interface based on profile described above. This radical personalization relies on the assumption that research preferences of a user cannot be fully logically or even probabilistically formalized (at most 0.01% of neurons in our brain work on logical, rational reasoning). Therefore, the interface should preserve and stress an intuitive character of the user choices, while nevertheless supporting her/his collaboration with the tools of ontological engineering. The model of PrOnto system assumes services and support to a research group of users (VCR) with functionalities serving an individual user or group collaboration. The model contains:

- *A radically personalized ontological model of the user*, composed of three layers as described above;
- *Document repository* **D**, containing documents interesting for the user or entire group of them (VRC) in the form of full text or a network link to such text;
- *A method of search and ranking of documents in the repository* for an individual user based on her/his radically personalized ontological profile (many methods are possible and the model of a user does not uniquely define such a method);
- *An agent of network search* (so called hermeneutic agent) that performs search in all accessible network – usually with help of accessible search engines – for new documents in order to enrich the repository, including a ranking method and(or) a decision rule;
- *Functionalities supporting an effective exchange of knowledge between users* that can enrich PrOnto system either for an individual or for group user. Such functionalities might include:
  - cataloguing documents for a group of users (VRC),
  - supporting research collaboration in the group (information about new documents judged as interesting by some users, etc.),
  - search for similarities in user interests, etc.

### 4.2. Searching for Information in Documents While Using Keywords

Documents in the repository must be indexed with respect to the keywords or keyword phrases. This is a standard problem known as *multiple pattern string matching*, searching for a pattern string (a keyword phrase) in a longer document. Because of large dimensions of documents and large number of pattern strings, it is important to select an algorithm with simplest, linear computational complexity; however, this complexity can be linear either with respect to the number of patterns strings (which can be very large), or, more advantageously, linear with respect to size of documents searched. An algorithm Aho-Corasick [30] was selected, implemented and tested, with the results shown in Fig. 1.

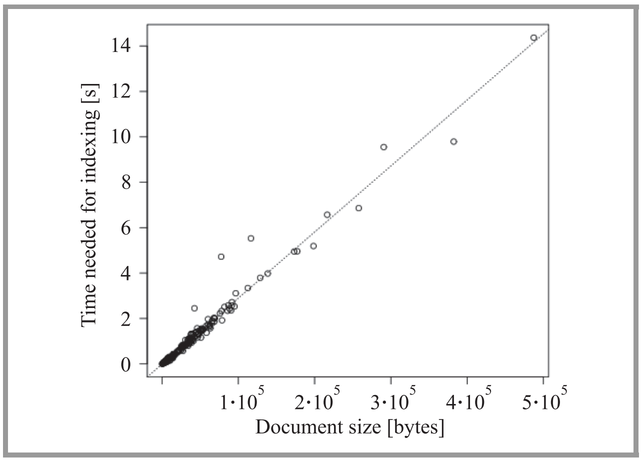



*Fig. 1.* Time needed for indexing as dependent on document size in bytes.

Another problem is a measure of importance of a document $d \in D$ with respect to a given key phrase $k \in K$. Initially, we selected the classical measure TF-IDF (*Term Frequency – Inverse Document Frequency*). The value of $\mathrm{TF-IDF}(k \in K,\ d \in D)$ grows proportionally to the frequency of occurrence of the phrase $k$ in the document $d$ and decreases inversely to the total number of documents containing $k$. We plan to investigate also other measures of importance, denoted here generally $g(d,k)$.

### 4.3. Importance of a Document with Respect to a Concept or a Set of Concepts

Another essential problem is a measure of importance of a document $d \in D$ with respect to a given concept $c \in \mathbf{C}$. If we have:

- set of documents $d \in \mathbf{D}$,
- set of concepts $c \in \mathbf{C}$,
- set of key phrases $k \in \mathbf{K}$,
- set of importance coefficients $f \in \mathbf{F}$ defining the relations between $c$ and $k$, a function $f : \mathbf{C}\mathrm{x}\mathbf{K} \to \mathbf{R}$,
- function $g : \mathbf{D}\mathrm{x}\mathbf{K} \to \mathbf{R}$ defining the results of indexing documents (importance of a document for a given key phrase),

then it is possible to define a measure of importance of a document $d \in D$ with respect to a given concept $c \in C$ as a function $h(d, c)$, e.g. as follows:

$$h(d, c) = \sum_{k \in K} f(c,k)g(d,k)$$

Other formulae as the above weighted sum can be also used, if we interpret differently the importance coefficients $f \in \mathbf{F}$

(as fuzzy logic membership values, or aspiration levels for multiple criteria ranking). We display this measure in the user interface.

However, a more important issue is the use of such measures in overall ranking of a set of documents with respect to entire personalized ontological profile, i.e., the entire set of concepts $\mathbf{C}$. A general way of defining a measure of importance of a document $d \in D$ to the entire profile (perspective, horizon) of the user is to treat each concept $c \in \mathbf{C}$ or, rather, each related measure $h(d,c)$ as a separate criterion of importance and then use methods of ranking related to multiple criteria decision making or to fuzzy logic; this will be the subject of further studies. A simple way is just to assume equal importance of each concept and just to sum measures $h(d,c)$ over $c \in \mathbf{C}$, or take a minimum of $h(d,c)$ over $c \in \mathbf{C}$ if each concept is considered essentially important.

### *4.4. Enriching Document Repository*

One of basic functions of PrOnto is to support user's including documents to enrich document repository. A special addition to the Firefox search engine was developed to support this functionality – see Fig. 2.

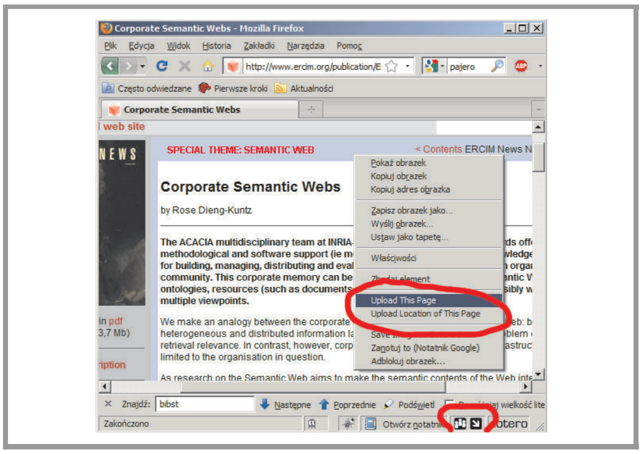



*Fig. 2.* Suggesting a WWW page for document repository, with marked elements of PrOnto Firefox Extension.

### *4.5. Multidimensional Search for Documents*

PrOnto system is equipped in an advanced search engine (concerning personal names, concepts, documents, key phrases), see Fig. 3, that presents the results of search in a multidimensional structure. The results of search for documents, based on a personalized ontological profile of one of the authors of this paper, are shown on the right side of Fig. 3. The concepts, shown on the left side, come from ontological profiles of many users, but the author of this profile selected those marked by ▶. When selecting a concept for more specific definition of importance coefficients $f$, this icon changes to ▼ (as at the concept *library*) and a set of keywords is displayed, with a simple interface to define subjective values of $f$. The keywords might come from the profiles of all users, or a set of key phrases originally defined by the specific user.

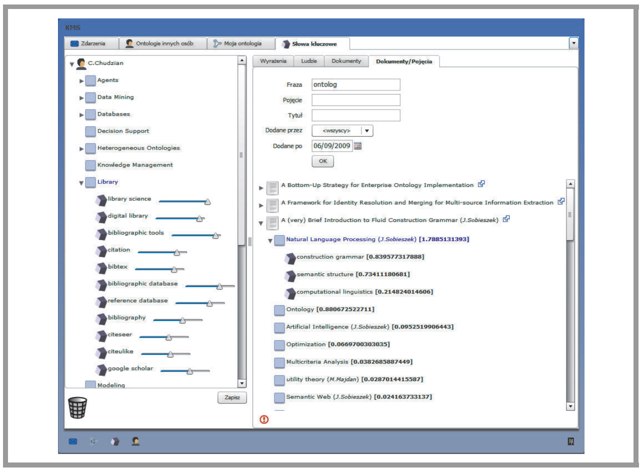


*Fig. 3.* Documents, concepts and key phrases.

### *4.6. Sharing Knowledge Using Ontological Models*

Problems of accumulating, organizing and sharing sources of knowledge are addressed in computer science for a long time. Recently, however, the interest in these problems is growing because of the importance of internet or intranet as a source of information and knowledge.

This trend has many forms: *social networks*, *communities of practice*, *peer to peer networks*, *virtual research communities*, etc. In these forms, ontological engineering tools are also used. For example, system OntoShare ([31] aims at supporting knowledge exchange in a community of practice, using a common ontology constructed for this community. Users are characterized by profiles selected from this common ontology (this is a difference from our approach: we *start* from individual profiles because we assume the sovereignty of the user). System checks similarity of profiles and suggests document sharing.

Another example is project SWAP (*Semantic Web and Peer-to-Peer*) [32], [33]. The main issue in this project is *Ontology Matching*, see [34]. Another product of this project

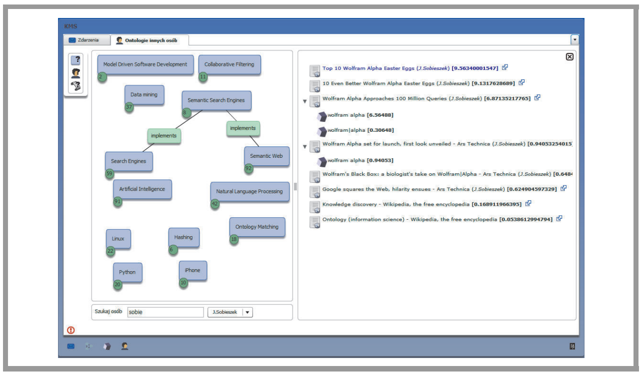


*Fig. 4.* Documents seen from a perspective of a given ontological profile.

is system Bibster [35] aiming at bibliographic information exchange in a distributed environment.
In the PrOnto system we assume that the users participating in a group (Virtual Research Community, VRC) approve sharing their personalized ontological profiles. Thus, one of functionalities of the system is to analyze importance of a document or a ranking of them *from another perspective* resulting from ontological profile of a different VCR member. This is shown in Fig. 4: on the left size a map of concepts is presented, on the left side a ranking list of documents, together with key phrases and corresponding values of $f(c,k)g(d,k)$.

### 4.7. Ontology Matching, Off-Line Analysis and Event Information

Another possibility offered by PrOnto is ontology matching. A user can see the concepts used in other ontologies than her/his own or even differences in relations between them. This is illustrated by Figs. 5 and 6.

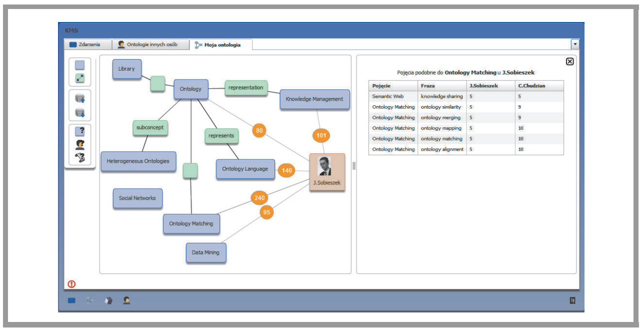


***Fig. 5.*** Similarity of user's profiles.

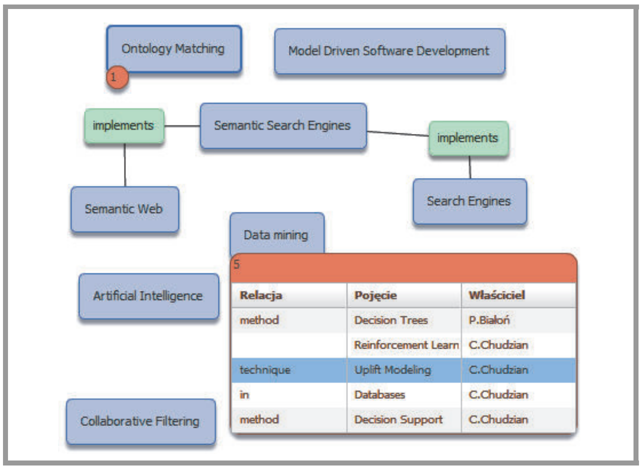



***Fig. 6.*** Checking differences in concept relations.

Beside interactive on-line work, PrOnto system performs also off-line analyses without user's participation. The results of such analyses are presented to users in the form of a list of events, such as occurrence of similar concepts in the profiles of other users, or enriching the document repository by new documents that might be interesting for a user.

### 4.8. Implementation Issues

PrOnto system was programmed using exclusively open source software. Some of such open source technologies used are already broadly applied, even included into commercial systems. We used a relational data base *PostgreSQL*, *Web Application Framework Django*, script language *Python* and the environment *Adobe Flex* for creating applications *Flash*. Moreover, PrOnto uses original codes written by authors in C language.

## 5. Conclusions

A prototype system PrOnto was developed in the *Requested Research Project "Teleinformatic Services and Networks of Next Generation – Technical, Applied and Market Aspects", Theme Group i: Systems Supporting Regulatory Decisions: Knowledge Mining in Large Telecommunication Data Sets*. This system realizes the perspective of *human centered computing* and is based on radically personalized ontological profiles of users that, on one hand, express intuition and tacit knowledge of a single user, but on the other hand enable an interaction with ontological engineering tools and with other users in a VRC.
There are many directions of future research on this system, see, e.g., [2]. Recently, these works were included into a new project SYNAT an we started to investigate diverse ways of ranking documents with respect to a personalized ontological profile with interpretations coming from fuzzy logic and multiple criteria decision theory.

## References

[1] C. Chudzian, E. Klimasara, J. Sobieszek, and A. P. Wierzbicki, "Wykrywanie wiedzy w dużych zbiorach danych: analiza tekstu i inżynieria ontologiczna. Sprawozdanie PBZ, *Usługi i sieci teleinformatyczne następnej generacji – aspekty techniczne, aplikacyjne i rynkowe, grupa tematyczna i: Systemy wspomagania decyzji regulacyjnych*: Warszawa: Instytut Łączności, 2009 (in Polish).
[2] C. Chudzian, J. Granat, E. Klimasara, J. Sobieszek, and A. P. Wierzbicki, "Wykrywanie wiedzy w dużych zbiorach danych: przykład personalizacji inżynierii ontologicznej", *Telekomunikacja i Techniki Informacyjne*, no. 1–2, 2011 (in Polish).
[3] I. Kant, *Kritik der reinen Vernunft*. 1781 (in Polish: I. Kant, *Krytyka czystego rozumu*. Warszawa: PWN, 1957).
[4] Z. Król, "The emergence of new concepts in science", in *Creative Environments*, A. P. Wierzbicki and Y. Nakamori, Eds. Springer, 2007.
[5] C. G. Jung, *Typy Psychologiczne*. Warszawa: Wydawnictwo KR, 2009 (in Polish).
[6] A. P. Wierzbicki and Y. Nakamori, *Creative Space: Models of Creative Processes for the Knowledge Civilization Age*. *Studies in Computational Intelligence*, vol. 10. Berlin: Springer, 2006.
[7] M. Polanyi, *The Tacit Dimension*. London: Routledge and Kegan, 1966.
[8] I. Nonaka, "The knowledge creating company", *Harvard Busin. Rev.*, vol. 69, pp. 96–104, 1991.
[9] I. Nonaka and H. Takeuchi, *The Knowledge-Creating Company. How Japanese Companies Create the Dynamics of Innovation*. New York: Oxford University Press, 1995.

[10] A. P. Wierzbicki and Y. Nakamori, Eds., *Creative Environments: Issues of Creativity Support for the Knowledge Civilization Age. Studies in Computational Intelligence*, vol. 59. Berlin: Springer, 2007.

[11] C. M. Bishop, *Pattern Recognition and Machine Learning*. Singapore: Springer, 2006.

[12] H.-G. Gadamer, *Warheit und Methode. Grundzüge einer philosophishen Hermeneutik*. Tübingen: J.B.C. Mohr (Siebeck), 1960.

[13] H. Ren, J. Tian, Y. Nakamori, and A. P. Wierzbicki, "Electronic support for knowledge creation in a research institute", *J. Sys. Sci. Sys. Engin.*, vol. 16, no. 2, 2007.

[14] H. Akkermans and J. Gordijn, "Ontology engineering, scientific method and the research agenda", in *Managing Knowledge in a World of Networks*, E. Motta, D. Sleeman, F. van Harmelen, V. Uren and A. Mille, Eds. Berlin-Heidelberg: Springer, 2006, pp. 112–125.

[15] M. Heidegger, *Sein und Zeit*. Halle: Niemayer, 1927.

[16] R. Mizoguchi, K. Kozaki, T. Sano, and Y. Kitamura, "Construction and deployment of a plant ontology", in *Knowledge Engineering and Knowledge Management*, R. Dieng and O. Corby, Eds. of Proc. 12th Int. Conf. EKAW 2000, Juan-les-Pin, France, 2000, pp. 113–128.

[17] O. Corcho, M. Fernández-López, and A. Gómez-Pérez, "Methodologies, tools and languages for building ontologies: where is their meeting point?", *Data & Knowl. Engin.*, vol. 46, pp. 41–64, 2003.

[18] H. S. Pinto and J. P. Martins, "Ontologies: How can They be Built?", *Knowl. Inform. Sys.*, vol. 6, pp. 441–464, 2004.

[19] E. P. Bontas and C. Tempich, "Ontology engineering: a reality check", in *The 5th International Conference on Ontologies, DataBases, and Applications of Semantics (ODBASE 2006)*, R. Meersman *et al.*, Eds., vol. 4275, *LNCS*. Montpellier: Springer, 2006, pp. 836–854.

[20] A. Gómez-Pérez, M. Fernández-López, and O. Corcho, *Ontological Engineering*. Springer, 2003.

[21] J. Curtis, D. Baxter, and J. Cabral, "On the application of the cyc ontology to word sense disambiguation", in *Proc. Nineteenth Int. FLAIRS Conf.*, Melbourne Beach, USA, 2006, pp. 652–657.

[22] A. P. Wierzbicki, "On the role of intuition in decision making and some ways of multicriteria aid of intuition", *Mult. Crit. Dec. Mak.*, vol. 6, pp. 65–78, 1997.

[23] A. P. Wierzbicki, "Intuicja z perspektywy technicznej: znaczenie zasady multimedialnej i zasady emergencji", *Wiedza a intuicja*, A. Motycka Ed. Warszawa: Wydawnictwo IFiS PAN, 2008, pp. 231–264 (in Polish).

[24] H. Ren, J. Tian, A. P. Wierzbicki, Y. Nakamori, and E. Klimasara, "Ontology construction and its applications in local research communities", in *Modelling and Decision Support for Network-based Services*, J. Granat and D. Dolk, Eds., 2008.

[25] J. Tian, A. P. Wierzbicki, H. Ren, and Y. Nakamori, "Constructing an ontology for a research program", *Int. J. Knowl. Sys. Sci.*, vol. 4, no. 2, pp. 59–67, 2007 .

[26] A. Schmidt and S. Braun, "Context-aware workplace learning support: concept, experiences and remaining challenges", in *Proc. First Eur. Conf. Technol.-Enh. Learning ECTEL 06*, Springer, 2006.

[27] C. Chudzian, "Ontology creation process in knowledge management support system for a research institute", *J. Telecommun Inform. Technol.*, no. 4, pp. 47–53, 2008.

[28] J. Sobieszek, "Towards a unified architecture of knowledge management system for a research institute", *J. Telecom. Inform. Technol.*, no. 4, pp. 54–59, 2008.

[29] D. Goldberg, D. Nichols, B. M. Oki, and D. Terry, "Using collaborative filtering to weave an information tapestry", *Communications of the ACM*, vol. 35, pp. 61–70, 1992.

[30] A. V. Aho and M. J. Corasick, "Efficient string matching: an aid to bibliographic search", *Communications of the ACM*, vol. 18, pp. 333–340, 1975.

[31] J. Davies, A. Duke, and Y. Sure, "Ontoshare – an ontology-based knowledge sharing system for virtual communities of practice", *J. Universal Comput. Sci.*, vol. 10, no. 3, pp. 262–283, 2004.

[32] M. Ehrig, P. Haase, B. Schnizler, S. Staab, C. Tempich, R. Siebes, and H. Stuckenschmidt, *SWAP: Semantic Web and Peer-to-Peer Project Deliverable 3.6 Refined Methods*, 2003 [Online]. Available: http://swap.semanticweb.org/public/Publications/swap-d3.6.pdf

[33] M. Ehrig, C. Tempich, and Z. Aleksovski, *SWAP: Semantic Web and Peer-to-Peer Project Deliverable 4.7 Final Tools*, 2004 [Online]. Available: http://swap.semanticweb.org/public/public/Publications/swap-d4.7.pdf

[34] M. Ehrig, *Ontology Alignment: Bridging the Semantic Gap (Semantic Web and Beyond)*. New York: Springer, 2006.

[35] J. Broekstra, M. Ehrig, P. Haase, F. Van Harmelen, M. Menken, P. Mika, B. Schnizler, and R. Siebes, "Bibster – a semantics-based bibliographic peer-to-peer system", in *Proc. Third Int. Semantic Web Con.*, Hiroshima, Japan, 2004, pp. 122–136.

---

**Cezary Chudzian** received his M.Sc. in computer science from the Warsaw University of Technology in 2002. He is a researcher at the National Institute of Telecommunications. Currently he works on his Ph.D. in the area of knowledge management. His main scientific interests include: practical applications of knowledge discovery techniques, machine learning theory, knowledge management, global optimization, and advanced software engineering.
E-mail: C.Chudzian@itl.waw.pl
National Institute of Telecommunications
Szachowa st 1
04-894 Warsaw, Poland

**Janusz Granat** received his M.Sc. in control engineering (1996) and his Ph.D. (1997) in computer science from the Warsaw University of Technology, Poland. He holds a position as an Assistant Professor at the Warsaw University of Technology, and is the leader of a research group on applications of decision support systems at the National Institute of Telecommunications in Warsaw. He lectured decision support systems and various subjects in computer science. His scientific interests include data mining, modeling and decision support systems, information systems for IT management. Since 1988 he has been cooperating with IIASA. He contributed to the development of decision support systems of DIDAS family and the ISAAP module for specifying user preferences. He has been involved in various projects related to data warehousing and data mining for telecommunica-

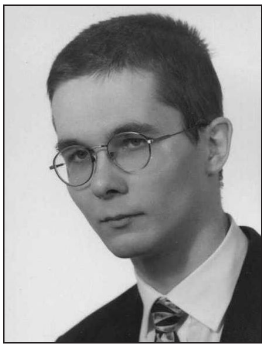
tion operators. He was also involved in EU MiningMart project.

E-mail: J.Granat@itl.waw.pl
National Institute of Telecommunications
Szachowa st 1
04-894 Warsaw, Poland

Institute of Control and Computation Engineering
Warsaw University Technology
Nowowiejska st 15/19
00-665 Warsaw, Poland

**Edward Klimasara** obtained his M.Sc. in Mathematics from Warsaw University, Poland in 1977. He is a senior specialist at National Institute of Telecommunications in Warsaw. Currently he takes part in projects: SIPS, EDFAS, SYNAT. His main scientific interest includes: knowledge management, IT systems for telecommunications and healthcare, network security, quality management, simulation techniques.

E-mail: E.Klimasara@itl.waw.pl
National Institute of Telecommunications
Szachowa st 1
04-894 Warsaw, Poland



**Jarosław Sobieszek** received his M.Sc. degree in computer science from Warsaw University of Technology, Poland, in 2002. Currently he is a researcher at National Institute of Telecommunications, where he prepares his Ph.D. thesis in the area of knowledge management. His research interests include machine learning, artificial intelligence, knowledge management and model-based approaches to software development.

E-mail: J.Sobieszek@itl.waw.pl
National Institute of Telecommunications
Szachowa st 1
04-894 Warsaw, Poland

**Andrzej Piotr Wierzbicki** got his M.Sc. in 1960 in telecommunications and control engineering, Ph.D. in 1964 in nonlinear dynamics in control, and D.Sc. in 1968 in optimization and decision science. He worked as the Dean of the Faculty of Electronics, Warsaw University of Technology (WUT), Poland (1975–1978); the Chairman of Systems and Decision Sciences Program of International Institute for Applied Systems Analysis in Laxenburg n. Vienna, Austria (1979–1984). He was elected a member of the State Committee for Scientific Research of Republic of Poland and the Chairman of its Commission of Applied Research (1991–1994). He was the Director General of the National Institute of Telecommunications in Warsaw (1996–2004). He worked as a research Professor at Japan Advanced Institute of Science and Technology (JAIST), Nomi, Ishikawa, Japan (2004–2007). Beside teaching and lecturing for over 45 years and promoting over 100 master's theses and 20 doctoral dissertations at WUT, he also lectured at doctoral studies at many Polish and international universities. Professor Wierzbicki is an author of over 200 publications, including 14 books, over 80 articles in scientific journals, over 100 papers at conferences; the author of 3 patents granted and industrially applied. His current interests include vector optimization, multiple criteria and game theory approaches, negotiation and decision support, issues of information society and knowledge civilization, rational evolutionary theory of intuition, theories of knowledge creation and management.

E-mail: A.Wierzbicki@itl.waw.pl
National Institute of Telecommunications
Szachowa st 1
04-894 Warsaw, Poland