

2013 AASRI Conference on Intelligent Systems and Control

The Supportive System for the Processing of Human and Non-Human Knowledge Sources and Combining Mining Techniques

Stefan Svetsky^a, Oliver Moravcik^{b*}

^a*Slovak University of Technology - Faculty of Materials Science and Technology in Trnava, Paulinska 16, 917 24 Trnava, Slovakia*

^b*Slovak University of Technology - Faculty of Materials Science and Technology in Trnava, Paulinska 16, 917 24 Trnava, Slovakia*

Abstract

In real life, an individual user must simultaneously process huge amounts of unstructured data, information and knowledge in various uncertain processes. Moreover, one must permanently switch between them. There are many single-purpose approaches for the processing of knowledge, mining techniques, soft computing or knowledge based systems. However, they are not very user friendly; this is why users need many software solutions to be sustainable. This paper refers to a multipurpose pre-programmed knowledge system called BIKE, which enables individuals to process knowledge and computerize knowledge-based processes as an all-in-one solution. The system works as an interactive software application based on database technology utilizing Boolean logic. It covers several fields in informatics, ongoing from an interdisciplinary definition of knowledge, knowledge tables, and the formulated batch knowledge-processing paradigm. This paper explains and illustrates how it works in practice on a personal level for individuals.

© 2013 The Authors. Published by Elsevier B.V. Open access under [CC BY-NC-ND license](#).
Selection and/or peer review under responsibility of American Applied Science Research Institute

Keywords: database technology; system for knowledge processing; mining techniques; technology-enhanced learning; FPW 2.6a

* Corresponding author. Tel.: +421 906 068165; fax: +421 33 5511758.
E-mail address: oliver.moravcik@stuba.sk

1. Introduction

When working with knowledge it is linked to the mental processes of the human mind. These processes are very complex and sophisticated, that is, they are automatically uncertain and have a large variability of potential solutions. This complicates any computer support related to the processing of data, information and knowledge. There are many, however, just single purpose approaches for the processing of knowledge or mining techniques, although an universal approach or software application is still missing.

Knowledge is a key element in teaching and learning, decision-making, knowledge management and other daily-performed processes by individuals. For instance, if one has perfect software for the acquisition of information, it does not automatically mean that it would be a suitable basis for the creation of teaching texts (what about the quality concerning educational content?). As well, when teaching, communication, feedback, what about quality of didactical approaches? Perhaps such software could be perfect for solving automation within a technical system. In living systems it is uncertain that the students or individual users (via informal learning, self - study) would understand the subject matter (it's a long journey with many processing steps, from the point of selecting optimal content from information sources until the transfer into the learner's brain). Challenges for computer support in the view of education on an institutional level are discussed exhaustively in [1] (regarding knowledge transfer from textbooks to global social memory).

All knowledge issues mentioned above were solved on an individual's level within implementation technology-enhanced learning (as one of the European Union research policy priorities) when teaching bachelors. This resulted into developing a supportive personalised system BIKE (Batch Information and Knowledge Editor). The core of this system forms a database application programmed within the conventional FPW 2.6a database platform. However, the application is adapted to present technology, that is, the focus is on interoperability with common internet browsers, Windows, freeware, open sources, office packages, including other database platforms (MySQL, IBM DB2). The use of BIKE enables users to combine informatics activities together, which would otherwise require dozens of single-purpose software applications. The BIKE approach for processing knowledge and mining techniques is illustrated more closely in this paper.

2. The processing of knowledge within the supportive system BIKE

The BIKE system solves the issue of processing knowledge based on the batch knowledge-processing paradigm, the paradigm that enables a user to work with a vast amount of knowledge in real time. The knowledge is processed in batches within the knowledge tables. It is natural for humans to work with a "vast amount" of "knowledge". Originally, the system was not developed with the ability to process knowledge, but was used as an aid to be sustainable in competitive market conditions. The method of how to process large amounts of *data* and *information* was first solved within an industrial R&D laboratory. Staff named the system Zapisnik (Writingpad). Afterwards in the university, this system was modified with *knowledge processing* by using BIKE within the implementation of technology-enhanced learning. The fundamental question of "what is knowledge" needed a solution in order to automate teaching processes. The problem was not only that any universal definition does not exist, but the approaches to knowledge are interdisciplinary different (e.g. see definitions in <http://oxforddictionaries.com/definition/english/knowledge>). In the BIKE system terminology, knowledge is defined as being *a set of information structured and unstructured, having a specified content stored in one row of the knowledge table within a default structure* [2]. This enables an individual to utilize the vast power of database technologies and existing programming languages. All outputs concerning the teaching of bachelors related to the progress in developing BIKE for knowledge processing

with a focus on the personalised approach and the automation of teaching processes, were continuously published, e.g. in [3-5].

One can imagine the system BIKE as a personal knowledge system or base which is empty at the beginning. Success in the processing of knowledge depends on the methods used for recording human or non-human knowledge into the knowledge tables. Figure 1. (a) illustrates examples of how content is stored in the knowledge table. It can consist from strings and texts, e.g. personal notes, ASCII characters (“non-human knowledge source”, e.g. a stored jpg- or mp3-file), or programming code which can be edited and launched directly from BIKE (here C++). This content in the memo-field is identified by data and information in other columns within one row. Figure 1. (b) illustrates a typical browsable output after the conversion of knowledge into the html – format. The converted content on the right window represents a scanned text from a scientific book; the left window displays the scrollable menu containing the knowledge.

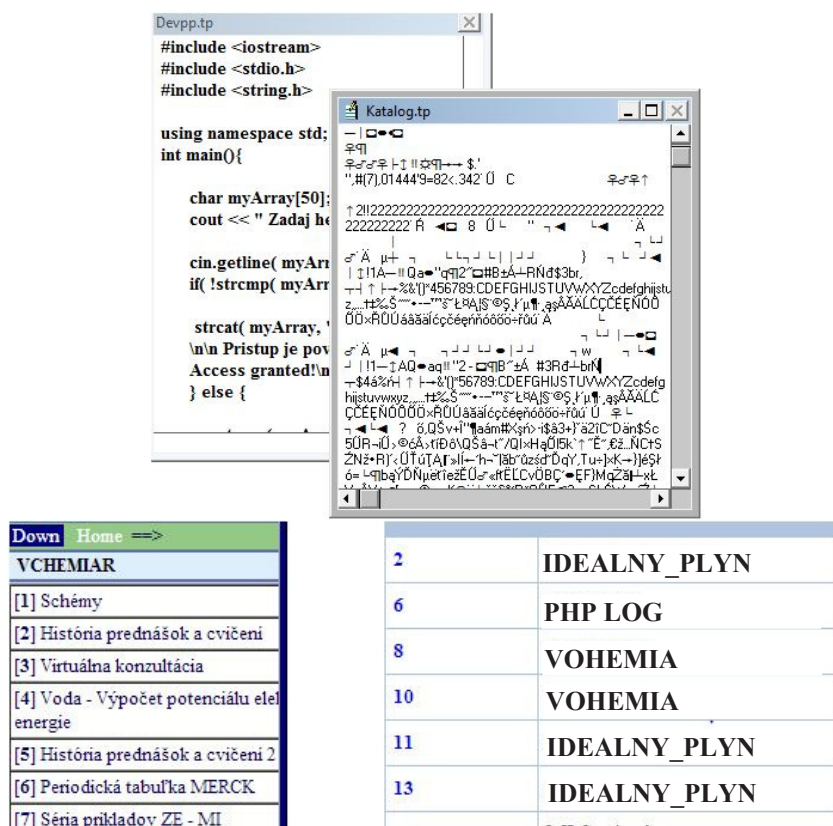


Fig. 1. (a) Examples of content categories in a row of the knowledge table; (b) Example of the browsable output (chemistry)

Figure 2. (a) illustrates a method of how people use and process knowledge when solving their daily tasks. Users process information from various sources (internet, personal know-how, journals, mobile phones etc.) step-by-step through many levels to transform it into knowledge which is in their heads. This knowledge is used to support personal processes; however, people must switch among them. The processing of knowledge

is cyclic, because a new knowledge output becomes an input for repeatable processing. Fig. 2. (b) illustrates how this works in teaching practice. In the previous years, BIKE built a supportive infrastructure that works as a *personalised technology-enhanced learning system*. One part of the system with knowledge tables is on the teacher's computer and the other part is placed on the faculty's server (teaching and study content, libraries, communication channels, calculation area, tests, etc.). This enables the teacher to solve *automation of knowledge processing* in order to make innovations and to design new teaching methods. Within participatory action research, it was found that content, teaching and communication processes must be unified in order to be computerized. Figure 2. (b) from [4] illustrates that if we have an implemented infrastructure and knowledge based system, we can start learning analytics and mining techniques.

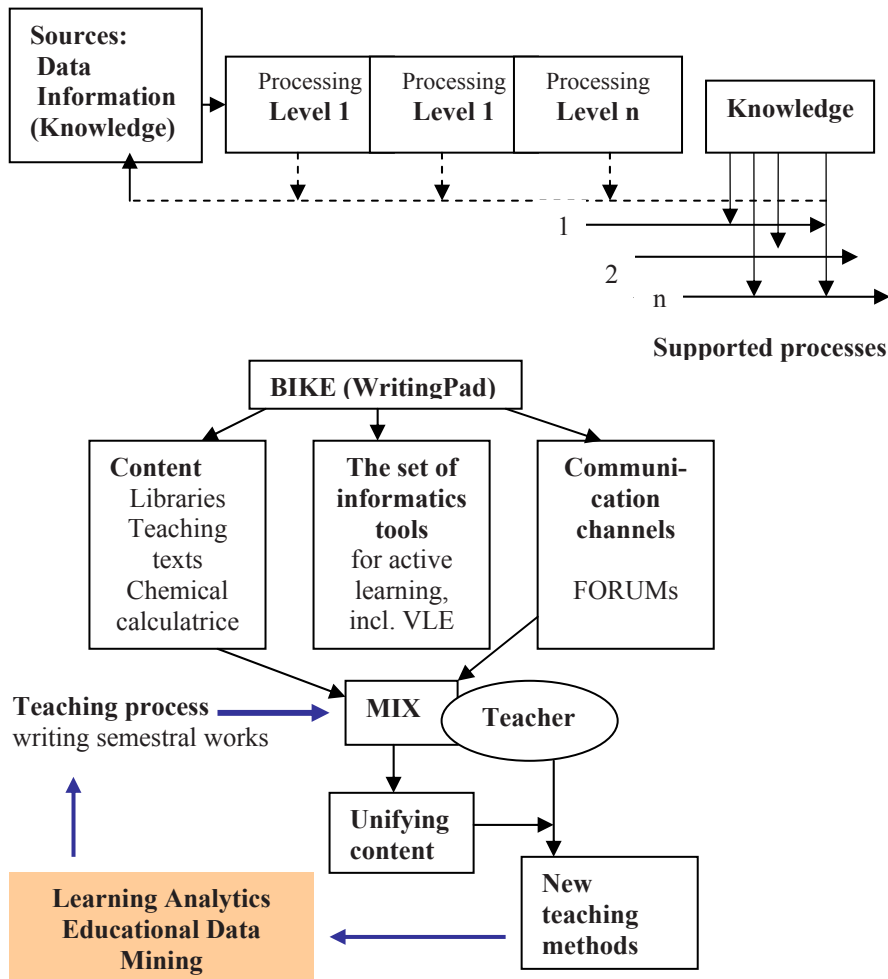


Fig. 2. (a) Scheme of knowledge processing by a human; (b) Scheme of implementing new teaching methods by BIKE

3. Examples from the solving of knowledge processing and combining mining techniques in practice

The knowledge system BIKE is useful when analysing documents, browsing or for advanced multi-language searches in internet databases on-line or off-line. This is programmed as “strings mining”. It does not matter what is mined (data, information, knowledge) and from where (WEB, personal client computer, CD-ROM, USB), e.g., the ICL conference proceedings were analysed (Villach, 2012). Around 200 pdf-files (papers) via USB were converted into one PDF-file and then to a text-file (by Nuance software). A new knowledge table was created and the text-file was appended into one row. Dividing a 6 MB text into single words would take hours therefore the blocks of texts (keywords and abstracts) were extracted into two tables, each with approximately 200 rows, these were then converted into a browsable format and single words. This resulted in a thousand keywords that were analysed in view of the highest frequency of appearance. Fig. 3. illustrates these steps on screenshots (knowledge tables -- browsable html-file -- outputs), as well the quantity of keywords that appear, e.g. six papers were focused on technology-enhanced learning (teacher’s expertise). Regarding the keyword *component* (see the 47 occurrences) the computer only indicated a string *component* from the *Keyword-component*. This is not a keyword!

Knowledge workers need to permanently search, browse, and mine knowledge from specialised internet databases focused on research news, technological data, scientific information, etc. Moreover, these are mostly multilingual, thus a cognitive load on individuals is extremely high, for example, the EU CORDIS Marketplace database focuses on results from funded projects (<http://cordis.europa.eu/marketplace/>) in five languages. Fig. 3 illustrates how the user can create a set of browsable navigation files for research purposes.

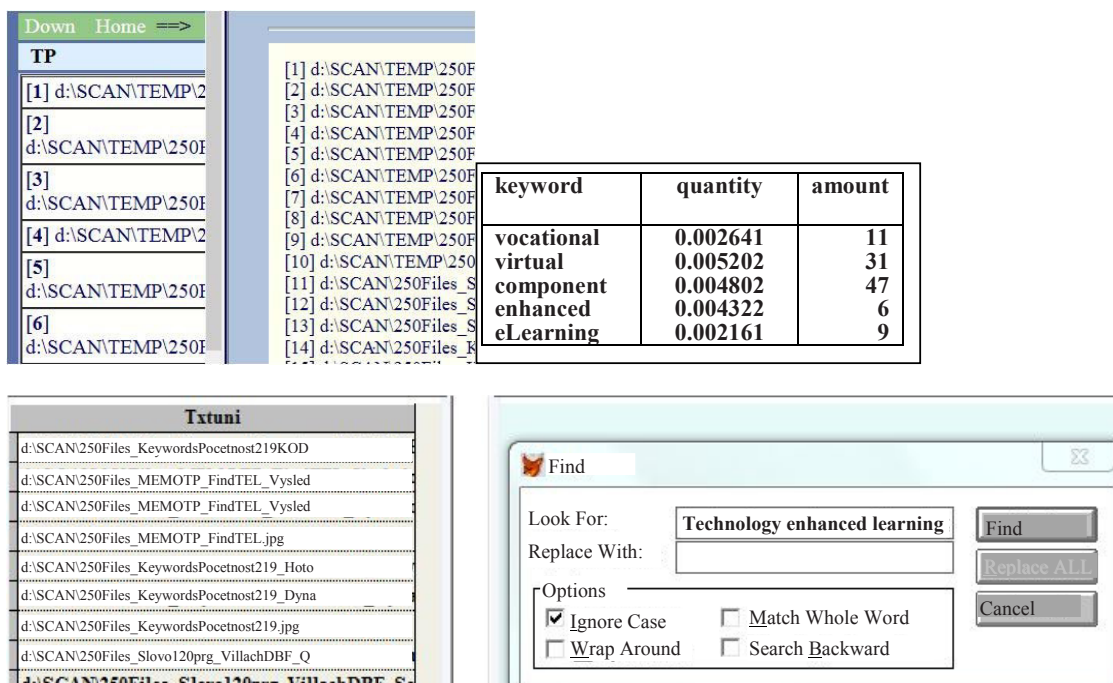


Fig. 3. Scheme of a set of screenshots from an analyse of documents

Datum	Txtuni
27.03.2013	10597: Marketplace
27.03.2013	10108: Marketplace
27.03.2013	10109: Marketplace
27.03.2013	10110: Marketplace
27.03.2013	10111: Marketplace
27.03.2013	10112: Marketplace
27.03.2013	10
27.03.2013	[3] 10595: Marketplace
27.03.2013	[4] 10594: Marketplace
27.03.2013	[5] 10593: Marketplace
27.03.2013	[6] 10592: Marketplace
27.03.2013	[7] 10591: Marketplace
27.03.2013	[8] 10590: Marketplace
27.03.2013	[9] 10589: Marketplace
27.03.2013	[10] 10588: Marketplace
27.03.2013	[11] 10587: Marketplace
27.03.2013	[12] 10586: Marketplace
27.03.2013	[13] 10585: Marketplace
27.03.2013	[14] 10584: Marketplace
27.03.2013	[15] 10583: Marketplace



Fig. 4. Browsing 10 157 records (offline / online) in Marketplace

4. The comparison of the BIKE system approach with state-of-the-art

The bike system was developed with the intention that a computer, like a user's partner, should expand one's quality of his life and sustainability. According to one of BIKE's authors' opinions (Svetsky) the paradigm and definition of knowledge is analogous to how people think and act. People perform processes that use knowledge that needs to be worked into daily life. Knowledge is also a key element within teaching and learning. The first stage of development was focused on the programming of knowledge flow (I) between information sources and knowledge tables (basically, content issues were computerised). The following came into question, what about knowledge flow (II) – between a knowledge base and teaching processes? What about knowledge transfer between teacher and students? The answers pointed to activities such as, the teacher must prepare teaching material, texts from the knowledge flow (I) and must communicate with students and concentrate knowledge (subject matter) in the classroom in a limited time and space whilst accounting for didactic quality. Thus, the knowledge flow (II) should also be programmed. This resulted in a strategy that technology-enhanced learning is to be understood as the *automation* of teachers or learning activities, where knowledge is the key parameter. Following this, it was found that writing codes for solving the knowledge flow (II) requires "helping the computer" by *unifying* the *content* and the *teaching process* in order to be *computerizable*. Without this approach, any programmer would not have an idea of how the communication and teaching process must be programmed. Therefore the approach for technology-enhanced learning must be

primarily *education - driven*. This aspect is often emphasised in actual literature, e.g. in [6]. Very precise information related to education - driven computer support was given in [7]. This trend is also visible in the European Union ICT policy stated in the FP7 research programme, e.g., previous target outcomes for Objective 8.1, technology-enhanced learning were: *technology-enhanced learning systems, educational technologies for science, technology and maths, computational tools, exploratory activities for fundamentally new forms of learning through ICT*. For instance, an expected impact was to “Unlock the potential of the individual by a stronger and smarter adaptation and personalization of educational technologies”. Thus, the developing supportive system of BIKE is in compliance with this research policy (this is illustrated in Fig. 2). The latest FP7 ICT calls are focused on *learning analytics* and *educational data mining*. In this context, further activities are also focused on mining techniques (also illustrated in Fig. 2). Concerning mining techniques, a knowledge worker does not differentiate between data- / text- / knowledge- / WeB mining. In compliance with the definition of knowledge, this is understood as *knowledge mining*, respectively *strings mining*. This enables one to apply repeatable programming codes. The compliance with *text mining* can be compared with [8, 9]. The intermediate forms for knowledge representation on the level of textual data, which were discussed in [9], evocate a new idea of understanding knowledge tables as intermediate forms on a higher level (BIKE’s authors operate with the knowledge table just as one does with a data type). BIKE’s approach is similar to the trends for minimising information overload, sharing information or solving multitasking, e.g. within the EU project Active it is argued that to switch between task contexts to improve productivity of knowledge workers is one of the hardest things for workers to do [10]. Many shared points are in the field of *advanced search* and *information retrieval*. One can select from items of BIKE’s user menu to support processes or tasks in which knowledge plays a key role. One can utilize the knowledge for “whatever” one wants - searching, extracting, mining, converting, exchanging, sharing, joining or making batch internet retrievals (the latter is used by students). All this occurs within the knowledge tables. Thus, BIKE’s applications are also directed at the field of *knowledge management* due to the system working as a knowledge based software application. A similarity can be found with activities of the Department of Cybernetics and Artificial Intelligence of the Faculty of Electrical Engineering and Informatics [11]. After BIKE’s author’s lecture at this faculty, the students categorised BIKE as a *knowledge base repository system*. From another point of view, the system was presented as a *personalised knowledge management system* [12].

5. Conclusion

The supportive knowledge system BIKE was presented by examples of knowledge processing based on the definition of the knowledge and batch knowledge processing paradigm. This system covers dozens of informatics categories, inter alia, advanced search, retrieval and “string mining techniques” (both off-line and online). This in-house developed database application enables individuals to have support of human decision-making, teaching and learning, knowledge management, administration and other actions in a user friendly way. BIKE works as a user’s partner, it has its own “social memory” that increases an individual’s performance and automates its personal activities. The approach for the solving of knowledge processing was explained in detail and briefly illustrated by the use of the analysed documents and the exploration of a multilingual internet database Marketplace. The BIKE approach was also compared with state-of-the-art. One of the future focuses should be on learning analytics and analysis of students’ behaviour via the use of the BIKE system.

Acknowledgements

This work was supported within the funded project KEGA No. 047STU-4/2012: The building online teaching room for education in the field of machining of complex shaped components.

References

- [1] Saljö, R. Digital tools and challenges to institutional traditions of learning: technologies, social memory and the performative nature of learning. *Journal of Computer Assisted Learning*. Vol.26, pp. 53–64; 2010.
- [2] Svetsky, et al. The Implementation of the Personalised Approach for Technology Enhanced Learning. *Lecture Notes in Engineering and Computer Science*. Volume 2186, Issue 1, 2010, pp. 321-323.
- [3] Svetsky, S. et al. The Particular Approach for Personalised Knowledge Processing. In: *Advances in Intelligent and Soft Computing*. Springer-Verlag Berlin Heidelberg, 2012, pp. 937-946.
- [4] Svetsky, et al. The Implementation of the Personalised Approach for Technology Enhanced Learning. *Lecture Notes in Engineering and Computer Science*. Volume I, 2012, pp. 290-296.
- [5] Svetsky, S. The practical aspect of knowledge construction and automation of teaching processes within technology-enhanced learning and eLearning. *Habilitation thesis*. Slovak University of Technology; 2012.
- [6] Derry, J. Technology - Enhanced Learning: A Question of Knowledge. *Journal of Philosophy of Education*, vol. 42, pp. 505–519; 2008.
- [7] Goodman Paul S., et al. *Technology Enhanced Learning: Opportunities for Change*. Laurence Erlbaum Associates, Mahwah, NJ, USA; 2002.
- [8] Drewes, B. Some Industrial Applications of Text Mining. *Knowledge Mining. Studies in Fuzziness and Soft Computing*. Springer - Verlag, Berlin Heidelberg. Volume 185, 2005, pp 233-238.
- [9] Justicia de la Torre, M. Consuelo, et al. *Text Knowledge Mining: An Approach to Text Mining*. Spain, ESTYLF 2008: XIV Congreso Español sobre Tecnologías y Lógica Fuzzy.
- [10] Putting knowledge in context to end information overload. *Research EU Results magazine*. IT and Telecommunication. No 11 April 2012.
- [11] Paralic, J., Paralic, M., Mach, M. *Support of Knowledge Management in Distributed Environment*. Informatica. Knowledge Based Software Engineering Information Technology. Vol. 25. Ljubljana; 2001.
- [12] Svetsky et. al. IT Support for Knowledge Management within R&D and Education. *ICL 2012. 15th International Conference on Interactive Collaborative Learning Villach, Austria*. - Piscataway : IEEE, 2012.