A Universal Paradigm for Integrating IT into Educational and Knowledge-Based Processes

Stefan Svetsky, Member, IAENG, Oliver Moravcik, Member, IAENG

Abstract—Despite the progress of ICT, the level of its integration into education is not satisfactorily resolved. A criticism is technocentrism, which results in the fact that teachers and researchers must adapt to the technology and not vice versa. In this context, multi-purpose software tools with longer life cycles are missing, as are personalized solutions where the technology is better adapted to teaching needs, and the user does not see any boundaries between off-line and online environments. In the framework of long-term research aimed at automatizing the knowledge-based processes, the authors have developed a specific paradigm based on the virtual knowledge unit (vKWU), which works as a knowledge representation. This specific data structure is controlled by own in-house software WPad (written by the main author of this paper) which enables batch processing, management and transfer of human knowledge that is inserted by the user into the vKWU in natural language. Because this vKWU is both human-readable and machine-readable, it allows the authors to solve the automation of knowledge-based processes and to model educational software as an all-in-one tool, and to simulate the individual's mental processes by using a personalized infrastructure (online/off-line). Such a solution eliminates the aforementioned shortcomings, thereby distinguishing it from the state-of-the-art, which requires combining several software in order to achieve the same output. The authors believe that their approach is beyond the state-of-the-art, because the educational content and processes are isomorphically connected with the technology by the associated virtual-human knowledge, and the WPad works as a black box. The paper describes the key milestones of their research and the current focus of their research. This is aimed at creating a combined infrastructure using clouds, virtual machines and off-line computers, and testing the transfer of knowledge for educational and collaborative applications within the V4+ACARDc project.

Index Terms—Cloud content management, computers and education, educational software, ICT for education, knowledge-based processes, knowledge representation

I. INTRODUCTION

Computer support for mental processes is a very complex problem because it involves the creation, processing, and management of human knowledge, and produces outputs for knowledge-based processes. These processes are very diverse and in terms of computer science are uncertain: they use unstructured data (information, content) stored in dozens of different formats. So, this also applies to educational processes, i.e., teaching, learning, formative assessment and associated processes such as research, publishing, self-study, which are also based on knowledge.

One should be aware that computers were invented for working with numbers and not with knowledge, which is a key element of the teaching processes. Therefore, if the computer is to process knowledge, it lacks a suitable knowledge representation, i.e., a certain structural assembly of zeroes and ones, as is needed for transmitting information. Information and knowledge in computer science are therefore something other than information and human knowledge in education, management, philosophy and other disciplines. As further pointed out by the authors, the specific problem is that the teaching processes are not sufficiently identified and described by the algorithms. For this reason, it is very difficult to integrate computer support into teaching by means of one software. As a result, teachers and researchers mostly use Office packages and additionally a set of single-purpose software, e.g., for processing images, graphs, audio and video applications. In other words, users primarily use existing software and technologies and they are trying to use it for IT support of the knowledge-based processes (teaching, learning, self-study).

Research related to integrating ICT into education is generally addressed within areas such as technology-enhanced learning and educational technology [1]-[3] – see also the authors’ publications in the previous WCECS and IAENG conferences (2010-2017), e.g. [4]-[5]. In the scientific literature, one can relatively often meet with the criticism that a technology-driven approach is prevailing, and the role of the teacher is secondary [6]-[7], or that the software design for technology-enhanced learning is inadequate [8].

In the present paper, the authors explain how they have dealt with these problems from the initial vision, through empirical research of personalized IT support to contemporary systematic research based on the universal representation of virtual knowledge that is both machine-readable and human-readable, and its transmission between users’ computers, WEB-space, virtual machines and clouds. The key component of the research is the development of a multi-purpose software, whose application user’s menu has currently hundreds of items (BIKE is installed on the computer of its designer and the WPad on the computers of the users).

Section II deals with the empirical phase of the research
under the umbrella of Technology-Enhanced Learning (TEL), and section III with the systematic phase of the research and the current research focus. In section IV, the essence and evolution of the paradigm of processing the human knowledge by individuals is discussed in a more complex way, in order to better understand the difference from the state-the-art in integrating IT into educational and knowledge-based processes. The novelty of the approach can be also seen on the basis of a registered utility model at the Patent Office related to processing the unstructured data (UV 7340/2016).

It should be emphasized that this is not a theoretical study, because all the outputs and applications have been already implemented in direct teaching and learning and IT support for individuals (research, language support, self-study); software application menus; knowledge tables and corpora, and into a virtual infrastructure (personalized virtual learning environment and space, virtual machines, clouds, WEB-hosting).

II. THE EMPIRICAL RESEARCH ON TEL

At the beginning of the research, the vision was to develop an all-in-one software system for generating, processing and managing general and expert knowledge at the individual level, because "to be sustainable requires for the individuals particularly to use personal expert knowledge base with the huge amount of steadily updated information (in a computer-readable and human-readable form)" [9]. Research started under the umbrella of TEL, which was a topic of the ICT calls within the 7th Framework Program of the European Union dedicated to the area of education, where the research team participated in the consortia KEPLER and L3Pulse submitting the project proposals. Research focus covered firstly topics such as the implementation of the personalized approach and educational-driven approach for TEL including the development of the informatics tools. For the purpose of IT support for teaching and learning, the single-user software BIKE/WPad was developed as a database application (it was written by the main author of the present paper), where the educational information was embedded into the classic database tables. A set of templates has been designed to allow information and knowledge to be bulk processed for automation of the educational processes as knowledge-based processes.

The principle of the construction and production of knowledge by WPad was described in [10], where the principle of the database paradigm for batch processing of information and knowledge was presented based on the outputs into the teaching of undergraduate students. They had WPad installed on their computers in the classroom. The IT support for several courses of study was carried out through the generation of the teaching materials by WPad (running off-line) and through communication channels (the PHP/MySQL application running on the faculty's server). Several theses can be also mentioned, in which students used the WPad software for generating small information systems for the field of occupational safety and health, cutting fluids, and ICT.

The result of the introductory phase of empirical research was that it developed a comprehensive system of IT support for undergraduate teaching based on the interconnection of the classroom computers (with the installed WPad) with the actual learning environment and space on the faculty's server. This system has proved to be very flexible not only in terms of students but also as a versatile tool for introducing new teaching methods, as shown in Fig. 1.

III. THE SYSTEMATIC RESEARCH ON INTEGRATING ICT FOR EDUCATION

Despite the fact that in the empirical phase of the research on TEL the students and teachers had at their disposal high-quality educational content (knowledge bases on the virtual learning space for several courses of study), functional technology (software and IT tools, tutorials) and suitable innovative teaching methods, the students were not willing to use it, either for learning or for self-study. It was deduced that the reason is very simple, i.e., to have content alone is not sufficient, because teaching and learning also consist of communication, knowledge transfer and educational processes in general, and this must be supported by computers.

Thus, even having high-quality content (knowledge bases, repositories, libraries, results of Internet retrievals) is not sufficient for teaching. There is a gap regarding the IT support of teaching and learning activities. Although a number of single-purpose software applications exist, no one of them is universal or has a suitable representation of knowledge. The state-of-the-art is based primarily rather on researching how existing market software can be fit for teaching. However, this does not guarantee an optimal solution and will mostly remain without a real impact on practice (this was the case with the failure of some academic projects in the 7th Framework Program, as stated by the Commission at the EDUCA conference in Berlin [12]).
The low interest of students in the voluntary use of the prepared high-quality teaching material has influenced the direction of further research. This was focused more on the IT-support of communication between the teacher and the students, as well as between the students and each other; the teaching process in the classroom with computers; and the transmission of the teaching material (content) between the classroom's computers and the faculty's server via communication channels. On this basis, new applications have been written, i.e., the user menu of WPad has been extended by dozens of new items. As an increase in the quality of teaching required to address the support of the collaborative learning activities as well as language support, some issues were solved within the international research consortia L3Pulse and PanEUlangNET, which participated within the ICT-calls for Learning Analytics (in the FP7 Framework Program of the European Union) and "Cracking the language barriers" (in the ICT 17-call of the Horizon 2020) [13].

A. Learning Analytics

The research on Learning Analytics (LA) already relied on the model of the vKWU and the paradigm of the batch information and knowledge processing driven by the WPad software. LA activities are in principle based on the collection of logs from the Internet (e.g., the number of students clicking per link in some time interval) and their statistical evaluation mainly for institutional needs. However, the presented vKWU interlinks the teaching content with the individuals directly through their personal cells (knowledge tables). Thus, for the purposes of LA, the human knowledge can be collected directly. In addition, the teacher's activities can be incorporated into LA by interlinking with his knowledge tables.

Fig. 2. illustrates this kind of personalized LA by explaining the bottom-up methodology, i.e., when the technology and the software are unfolding directly from the educational knowledge flow.

![Diagram of Learning Analytics](image)

Fig. 2. A comparison of the LA methodology with the state-of-the-art

But the state-of-the-art of LA is based on the top-down methodology, i.e., when the analytics is unfolding from data which identify the activities of students related to their teaching but not to the educational content (knowledge).

B. Human Language Technologies

Within the testing of the language support of teaching the undergraduate students, the use of Human Language Technologies was tested. The Text-To-Speech audio method proved to be progressive, also for minimizing plagiarism when writing semester works. The Speech recognition method was also tested but only by the author of the WPad (for Internet retrieval). Here, it should be added that at that time these technologies were not freely available as they are now. At that time, the idea of multilingual corpora was born, which would be built together by engineers and computer linguists. A CSA project was proposed by the authors within the ICT17-call of the Horizon 2020. In the context of that paper, it was planned to supply the engineering texts to the public portal, which would have been a source for automatic machine translation. These texts would have been generated by the WPad.

C. Computer Supported Collaborative learning (CSCL)

For elimination of the lack of knowledge of the chemistry of students coming to the university, the WPad retrieving function was used and every student had to process some theme from the foundations of chemistry, e.g., oxidation, reduction, chemical laws, and input it to the WPad-tables on his classroom's computer. By collecting the content from all these computers, the collaborative teaching material for the basics of chemistry was created. It was then used by the students in the next years. A similar methodology was also used in other courses of study. The greatest effectiveness of collaborative learning was achieved when teaching programming languages. In addition, in the case of teaching the C++ language, WPad served also as a source code editor. The development of the CSCL resulted in a significant finding, e.g., that CSCL cannot be optimally solved if the teaching algorithms are not identified and described. Namely, if these are not known firstly, it is not possible to write applications (informatics algorithms). This finding may appear to be normal, but this is not described in the CSCL literature (similarly, that the life cycle of software and hardware is shorter than required by long-term teaching, as highlighted by the authors in [14]). Moreover, a third condition for successful CSCL design is writing source codes that ensure compatibility with the operating system (Windows) and the online infrastructure. The issue of CSCL was presented by the authors for example in [14]-[16].

D. Some aspects of the systematic research approach

Despite the successful implementation of IT support for the teaching of undergraduate students (around two thousand for around ten years), the authors could not find a specific scientific journal for publishing their model of vKWU. Compliance was finally found in the cybernetics literature [17]-[18]. On this basis, one can imagine vKWU as a "switch" between the mental processes of the user and the physical processes of the machine. Fig. 3 illustrates this isomorphic relationship, so in this view, WPad works as a black box. The current research of the authors is therefore focused more closely on the "phenomenon" of virtual knowledge (vKWU) and its application to information and knowledge management.
Within the V4+ACARDC project, a transfer of knowledge (knowledge tables) and content (files) between the computers of the researchers and online spaces (clouds, virtual machines, WEB-hosting) is modeled. This is necessary for the design of the educational packages and modeling the support of writing. In the future, there is also the intention to transfer the existing virtual learning space to clouds and virtual machines. It can be mentioned additionally that the vKWU model was tested as an alternative way of metadata maintenance in collaboration with the authors of [19].

It should be also explained that there is a difference between using virtual machines and the BOX cloud service. The virtual machine is basically a remote desktop with the Windows operating system where WPad is installed and shared together by all partners of the project for exchanging the knowledge tables and documents (files). Such client-server exchange of tables, into which tacit knowledge is inserted in natural language, has not been described in the literature. Namely, it is common among computer users to transfer content using the standard file formats, e.g., PDF, DOC, TXT, HTML, PNG, MP3. The latest trend is to use a cloud-based service for these purposes.

In this case, WPAdV4 works as a client-server solution, so it enables users to transfer both knowledge tables and computer files between the offline and online environments. In combination with the BOX, due to the possibility of synchronizing (automatic transfer), the designer can browse the BOX cloud content from his computer, mobile and project partners from their notebooks. Fig. 4 illustrates such transfer among the project partners from the Czech Republic, Poland, the Slovak Republic, Hungary, and Ukraine.

**Fig. 3. Scheme of functioning of the in-house software as a black box [5] (owned by the authors)**

**Fig. 4. Infrastructure of the BOX-cloud content management**

**IV. EVOLUTION OF DESIGN OF THE PARADIGM FOR EDUCATIONAL KNOWLEDGE PROCESSING**

The state-of-the-art in the field of integrating IT into education is characterized by using the existing technology in the teaching processes, or single-purpose software, rather than to consider teaching processes as knowledge-based, where the educational knowledge is a controlled process-parameter. After a detailed survey of the scientific and patent literature aimed at ICT support for education, to the best of the authors’ knowledge, no similar research approach was found that would be based on a representation of knowledge that is both machine-readable and human-readable even in the users’ natural language. On the contrary, the existing solutions are based mostly on using machine languages and schemas (e.g., RDF, OWL).

On this basis, a specific paradigm has gradually evolved when teaching undergraduates and carrying out academic research. In this case, an individual user works with his computer in a different way from what is common. In short, as a calculator is used to work with numbers, the WPad software is used to work with human knowledge.

**A. Research approach**

To better understand this difference, this section describes a development of this research approach when solving the ICT-support for education which led to the creation of a new paradigm of the processing of human knowledge by computers. This description is needed because, in the experience of the authors, for a third person it is very hard to understand how WPad, as an all-in-one, multi-purpose solution, can be based on such a simple principle as virtual knowledge. While WPad works as an empty knowledge base system, into which users can insert the educational content and their tacit knowledge, most users are used to working with the Office packages.

At the beginning of the authors’ research (2006-2008), there was the vision that knowledge workers (researchers, teachers) should be equipped with IT tools for processing knowledge at the same level as the "contemporary soldier" is equipped with technological guns. In other words, to equip them with a universal "all-in-one" software that would allow them to process a large amount of information and
knowledge (“personal big data”), and to produce outputs into educational processes operatively. At this stage, the empirical research under the umbrella of TEL was started with its focus on the technology-driven approach. A development of the WPad software and the associated paradigm of batch processing information and knowledge has enabled processing a huge amount of human knowledge at the level of the individual for constructing the educational content (without, at that time, the idea of the vKWU). Within the design and testing of the applications to teaching, the research approach was already presented as education-driven, where the key player of the teaching processes is the teacher, e.g., in [10]-[11].

B. Knowledge representation design

After mastering the mass creation of educational content, it became clear that it is necessary to also solve the IT support of the communication and transfer of the educational content between the teacher and students within the teaching processes. Although a creation of the content base and knowledge repositories is sufficiently covered by the software solutions, there is still a lack of software solutions simulating educational activities and processes (and so, not covering the educational content).

As for this software gap, the absence of algorithms describing educational activities was identified by the authors. This fact does not allow a designer to write informatics algorithms for teaching applications. From the point of view of the automation of knowledge-based processes, where knowledge is a key parameter, it is unknown "what" knowledge is and "how" to define it. The model of the above-mentioned representation of knowledge (vKWU) solved this issue because it fulfills the principles of both Computer science and education, as well as being understandable by a layman. In this context, WPad worked as the educational software/technology within the teaching of several undergraduate courses of study. Over two thousand undergraduate students have used it for up to several years, and a dozen diploma theses were solved for the design of “small” information systems.

At that time, the authors tried to analyze why their approach functioned universally and why the WPad enables users (teacher, students, researchers) to solve education cases, for which solution some software products would be otherwise required. The result was that the key point would be that WPad uses a default structure that simulates and represents any human knowledge in an optimally way. Within the specific data structure, e.g., in the form of the classic database table, a block of metadata can be fixed as human information, and the content of the human knowledge can be fixed in a block, in which the larger ASCII texts can be embedded/inserted (note that if the block was empty, the knowledge table record would be considered as primitive information).

Similarly, a layman has any event or activity in his brain fixed as a mix of information (metadata) and content (human knowledge). This is in compliance with the hierarchy: data-information-knowledge (e.g., in the field of knowledge management). Therefore, the authors named the default structure “the virtual knowledge unit” (vKWU).

The universality of this model of knowledge representation is due to the fact that it is in compliance with the opposite hierarchy of information and knowledge, as in Computer Science. In the ICT field, there is a basic unit of information: a "bit" with values 0 or 1 (so, not just any data). In other words, computers understand information through the data types and structures (as a default set of information consisting of zeroes and ones). The computer thus reads the vKWU-data structure as information, but a user reads it as human knowledge (content) in natural language. As was mentioned above, the authors found finally a match of their research approach in the cybernetics literature, where the teaching process was described as a cybernetic system with feedback. In view of this, WPad provides an isomorphic interconnection between mental processes and machine processes, so the computer can process the virtual knowledge as information immediately and much more quickly than a human being.

The following systematic research already uses the "phenomenon" of the virtual unit of knowledge (vKWU) as an empty primitive intelligent structure. By filling this structure with knowledge, a personal knowledge base is created. In addition, WPad also works as a text corpus, as a Knowledge Management System [16]. On this principle, it is possible to process different categories of educational domains, which content is processed in WPad, as illustrated in Fig. 5 (WPadV4 is an update of WPad).

The authors gradually expanded their research focus on virtual machines when they began testing text corpora in a virtual environment. In addition, the WPad software designer had his personal remote desktop with Windows 7, where he tested a client-server application. Another virtual machine, which was shared by several educators, researchers, and PhD students, was then introduced and modeled for various CSCL applications. WPad has also been tested as a Knowledge Management System and to convert tacit knowledge to explicit knowledge [16].
C. The principle of integrating IT into education

In this phase, the authors' research has come into a new dimension after finding that if we have "one" educational knowledge, which is stored in a text corpus, we can use it for "many" teaching and learning activities, because they consist of knowledge flow running over time. This seemingly primitive finding has a significant impact on the potential automation of educational activities as the typical knowledge-based processes. "Many" activities - which means "many" of teaching algorithms and "many" informatics algorithms means that "many" source-codes must be written for a successful integrating IT into educational processes [20].

In this context, the authors analyzed three basic blocks which should be designed/solved synchronously:

1) Identification of teaching algorithms including related educational processes.
2) Assignment of information algorithms to the educational algorithms (design of applications).
3) Adaptation to the operating system, hardware, clouds, virtual machines, including compatibility with current software and heterogeneous formats.

Moreover, as for programming the teaching application, this requires not only selecting any knowledge content from a corpus or repository but also solving the problem of how to simulate its transfer within the communication of teacher and students (e.g., within a lecture, exercises), and logically its transmission between off-line and online environment. This all becomes possible to be solved by using the model of the virtual knowledge unit and represents a background for future research. It must be also mentioned that there are other barriers regarding the computerization of the teaching processes, including the short life cycle of software and hardware, which were discussed by the authors in [20]-[22].

D. The focus of the V4+ACARDC project

The current focus of the V4+ACARDC project is on the design of educational packages, so a modification of the WPAd for this purpose is needed (to WPAdV4). It was found that this demands an atypical way of working at a computer, which is not natural for single users. In fact, the design of the educational packages moreover requires working with folders and large numbers of files. Thus, additional source codes must be written that are typical for the functions of file managers. In these circumstances, the alignment of the three blocks mentioned above is essential (teaching algorithms – informatics algorithms – adaption to the operating system, software, and IT-infrastructure).

But in order to do this, a single user should also adapt to the computer's operation. In other words, developing a new paradigm for working with a computer is needed, which would also cover structuring the virtual knowledge, bulk transmitting of the knowledge cells, including transmitting the folders and files, even designing a tailor-made ontology and semantics. In this case, the alignment of the person with the computer depends on such details as using the appropriate file names and "natural text marking" (a method tested within the development of WPad). None of these steps are natural for people because they come with a great increase in entropy.

Within the project, WPAd is installed at the remote desktop with Windows 10 (virtual machine) and shared by all project partners for the collaborative design of the educational packages, which are created on the basis of the knowledge tables. As the files' repository, the project partners use the cloud service BOX, which enables Cloud Content Management. However, WPAd cannot be installed on the BOX, because the service has no operating system. On the other hand, the BOX enables the synchronized transmission of files from off-line to online and back, which is used to automatically transfer both the knowledge tables and the files directly by using the WPAd menu. Such a collaboration between the WPad and the BOX enables one also to solve some atypical personalized cases. For example, transferring the virtual knowledge and files between two notebooks by writing this paper, as illustrated in Fig. 6.

Fig. 6. Scheme of a collaboration of the WPAd with the Box cloud

V. CONCLUSION

Educational content (covering the domain ontology and semantics) and educational processes (teaching and learning and associated processes) are interconnected by human knowledge, which is a part of mental processes in general. On the other hand, there is a technology, i.e., an infrastructure in which informatics (computer) processes are running. A great challenge is that educational and informatics processes run commonly in isolation, or they are handled separately (e.g., collaborative educational algorithms versus technological algorithms), thus, they are not interlinked with knowledge. From a cybernetic point of view, the authors have solved this problem by simulating mental and educational processes with the aid of the design of a virtual knowledge unit (vKWU) which isomorphically interlinks the knowledge-based human processes and computer processes. Thanks to this universal knowledge representation, a multi-purpose software WPad has been developing and many outputs into teaching and research have been implemented that could otherwise only have been achieved by using dozens of existing software packages. The most important finding of this research is that teaching algorithms, informatics algorithms and adaption to technology must be solved synchronously. This is a challenge for future work of the authors.

REFERENCES


