COLLABORATIVE SYSTEMS FOR KNOWLEDGE ACQUISITION FROM SEMANTICALLY HETEROGENEOUS INFORMATION SOURCES

Ioana Andreea STĂNESCU *
Antoniu STEFAN **

Abstract: Novel advances in computing and communication technologies together with the rapid proliferation of information sources present unprecedented opportunities for the development of integrative learning environments. This article explores collaborative approaches to knowledge acquisition from distributed, autonomous, semantically heterogeneous data and knowledge sources in the perspective of building optimal blends of learning components.

Keywords: collaborative systems, knowledge acquisition, knowledge interoperability, SOA, mobile technologies

I. INTRODUCTION

People live, study and work in a semantically heterogeneous space that comprises rich, yet different information sources and applications, which are difficult to use at their full potential. While diversity is recommended as it sustains innovation and development, at the same time it represents a challenge for education and business organisations in terms of integration and exchange.

The last decade has registered progress towards building collaborative environments and systems that facilitate sharing and compatibility, and concurrently supports individuality and originality. Following the famous motto of unity in diversity, that sustains co-operation between different groups of people, in a single society, this paper aims to transfer this concept to an extended view that applies not only to social communication, but also to information and data applications, in a single information society.

The authors present this extended framework of knowledge acquisition in interoperable environments as a mean to promote collaboration between social actors and to assure access to semantically heterogeneous information sources for organisations.

The rapid development of high throughput data acquisition technologies have resulted in an explosive growth in the number, size, and diversity of potentially useful information sources, creating unprecedented opportunities for data-driven knowledge acquisition and decision-making in a number of emerging and increasing data-rich application domains, such as enterprise informatics or learning informatics.

The massive size, semantic heterogeneity, autonomy, and distributed nature of the data repositories present significant hurdles in acquiring useful knowledge from the available data. Against this background, there is an urgent need for software systems for collaborative knowledge acquisition from autonomous, semantically heterogeneous, distributed information sources [1].

This article presents a broad perspective on knowledge acquisition, as to include appliance not only in learning-oriented institutions, but also in business-driven organizations, as learning occurs both in formal and informal environments and sustainable knowledge acquisition is mandatory to support and improve performance, value-adding and competitive advantages.

Knowledge acquisition is one of those things that is easy to do badly and difficult to do well. At the heart of knowledge acquisition is the creation of a box of knowledge, which can be referred to
as a knowledge base, a knowledge store, a knowledge repository or an ontology. To create the knowledge base, or k-base for short, requires capturing knowledge from people with a great deal of expertise in a specific domain.

A definition of knowledge would be that it is equivalent to expertise, and expertise is the ability of people to do effective and efficient work, to deal with and solve complex issues creatively [2], and to take advantage of opportunities by making the most appropriate decision in particular situations. This represents the core of all learning activities, whether they occur in formal or informal settings, in specialised institutions or business organisations. To support this perspective in the context of the developments in computing and communication technologies, and of multi-source, differentiate data, it is necessary to address the issues of developing collaborative systems for knowledge acquisition.

This article explores collaborative approaches to knowledge acquisition from distributed, autonomous, semantically heterogeneous data and knowledge sources in the perspective of building optimal blends of formal and informal learning components for educational and business organization. The authors emphasis on the importance of knowledge interoperability as a sustainable practice for collaborative system development and present the benefits that advances in mobile technologies bring to our activities, sustaining enriched collaboration.

II. COLLABORATIVE SYSTEMS FOR KNOWLEDGE ACQUISITION

2.1 Knowledge acquisition: uses, benefits and challenges

Among the many definitions of knowledge that have been given along years, the authors resume to consider the following [2,3]:

- a highly-structured form of information;
- a kind of machine learning;
- a computer assisted technique;
- what is needed to think like an expert;
- what separates experts from non-experts;
- what is required to perform complex tasks.

Without proving all-inclusive, the following definition comprises a wider concept of knowledge, which allows the authors to establish broader settings for this article:

| Knowledge is the ability to manipulate data to perform skilfully |
| expertise | skill to transform information to make decisions |
| create ideas | solve problems |

As many artificial intelligence (AI) authors pointed out in the last two decades, knowledge acquisition is the key and bottleneck in expert systems development [3]. Scientific issues central to artificial intelligence underlie knowledge engineering in order to be enumerated as the parts of an expert system. When it comes to knowledge, it is important to consider: first, the problem of knowledge representation; secondly, the problem of knowledge utilization, and thirdly and most important, the question of knowledge acquisition [4].

Knowledge based systems acquire different types of knowledge. The reason people access knowledge in educational or work settings is to obtain information on what is or what is not a certain notion, or on what to do or how to a certain thing. In this regard, the most important types of knowledge that stand out are the declarative knowledge concerning facts, concepts and such, and the procedural knowledge concerning methods, procedures, rules, etc.

Knowledge acquisition is the activity of capturing expertise from people and other sources of knowledge to be used to help an organisation in some specified ways [3]. Performing knowledge acquisition requires expertise, and the knowledge engineers are the skilled people to do this.

As there is no point to knowledge acquisition unless we do something with the knowledge we have acquired, this paper includes some of the many uses for a k-base:

- Sharing knowledge with people: this would normally require the transformation of the k-base into a special website called a knowledge web. Usually, this is uploaded to the organization’s intranet system, so people can access the knowledge when they need it.
- **Sharing knowledge with computer systems:** this would normally imply re-coding of the k-base into the special format of an ontology that is embedded within a web system, so that software systems and services can access the knowledge when they need it.

- **As part of the development of an intelligent computer system:** this would imply re-writing the k-base as a knowledge document. This would be passed onto the organisation’s software developers so they can develop the system, which could be an expert system, a knowledge-based system or a knowledge-based engineering system.

Performing a knowledge acquisition project requires the creation of a k-base for one of these uses. However, it could be re-used for another purpose, such as providing material for a process redesign, or for the auto-generation of computer code. This is an important advantage of a k-base, that it can be used and re-used for different purposes.

The knowledge web, the ontologies or the knowledge documents specified above present the following main benefits:

- Knowledge web can be used to teach people who are just starting in an area and accelerate their progress up the learning curve. This use of knowledge acquisition can be accelerated further if the new starter acts as the knowledge engineer of the project (after a suitable period of training in knowledge acquisition).

- Knowledge web can be used to spread knowledge across the functional boundaries of an organization such as from design to manufacturing (and vice versa), or from technical people to financial people (and vice versa).

- Knowledge web can be used to archive knowledge for future generations. For example, it can store the reasons behind decisions that are made during the development of a new product. Some complex product such as military and aerospace products have an active life of many decades so it is vital that knowledge is passed down the generations in a format that is usable over a long span of time.

- Knowledge web can be used to reduce risks involved in losing access to people who have very specific knowledge. Thus, it is useful to create a knowledge web when people are close to retirement or when there only one or two experts in a domain.

- Ontologies can be used to provide a common language between software systems within a network. In this way, IT systems can use ontologies when reading, manipulating and using knowledge to perform all manners of tasks.

- Ontologies can be used to give structure to informal sources of data and information, such as Wikis and Blogs. In this way information can be fused and filtered to provide people with better views of the information they require.

- Expert systems can provide people with advice by replacing part of the reasoning that is performed by experts. In fact, experts can use such a system themselves to reduce workload when there is too much to do and too little time.

- Expert systems can enable inexperienced people to perform complex activities by providing suggestions and advice.

- Knowledge-based engineering systems can perform certain knowledge-intensive tasks such as designing complex components in a fraction of the time it takes a human to do the tasks.

All these examples prove that knowledge acquisition plays an important role in a number of fields, such as knowledge management, knowledge engineering, knowledge-based engineering and ontological engineering.

Knowledge acquisition implies certain difficulties [2] that developers need to take into account:

1. learning a complete unfamiliar domain when building a new knowledge base.
2. finding one or more experts who are experienced and are willing to cooperate.
3. extract as much knowledge as possible from the expert’s memory and behaviour.
4. what the expert provides is only raw material, often mixed with personal biases, even wrong conclusions; this imply the need to screen out, test and reorganize the knowledge obtained from the expert.
5. knowledge is not equal to experience; experience is not always representable; it may be fuzzy and inconsistent; it may appear in the form of inspiration and
randomly emerging ideas; the expert has difficulty in explaining it and the knowledge engineer has difficulty in understanding it.

6. there is no clear border between domain knowledge and common sense knowledge; the latter is informal, infinite, continuous and exists everywhere; it is difficult to decide what should be acquired and what should not be acquires.

7. knowledge cannot be acquired at one stroke; it has to be accumulated during a long process; even the most experienced expert is not able to provide this knowledge at a stretch.

Interoperability addresses some of these difficulties, enriching access to electronic knowledge data bases that already encapsulate expert data in various domains. In this regard, it is important to develop interoperable knowledge sources that permit cost and time reducing and better performance.

2.2 Knowledge sharing and interoperability

Knowledge processes can refer to tacit knowledge of experts combined explicit knowledge to create new products and services, or to innovation processes where people combine knowledge from different domains to create new knowledge. In terms of applications, we have distributed software project teams, design teams, planning and evaluation teams, client support teams, as well as the need for meetings during various stages of the business process. For example, producing a new car model can require well over 200 designers of different components to coordinate their activities and share their knowledge [5]. It is obvious that knowledge processes require intense exchange of information between team members to reach the desired outcome, and must go beyond simple exchange of messages or documents but operate interactively in a shared context.

A successful collaboration requires the combination of the following dimensions: the social culture, the ways to manage organizational knowledge and technology. The social environment is where people develop relationships, educational and work practices, which prove mandatory for sharing and creating knowledge in mutually acceptable ways for all the participating actors. Knowledge management provides support for interpreting information in its context and for distributing these interpretations. Technology

Knowledge management (KM) continues to evolve, and today it means many things to the myriad organizations that institute this paradigm. One thing to consider it that the practice of KM has its roots in a variety of disciplines, which include:

- **Cognitive science**: the study of the mind and intelligence, which comprises many disciplines including philosophy, psychology, and artificial intelligence (AI). Information learned from this discipline will improve tools and techniques in gathering and transferring knowledge.
- **Expert systems, Ai, knowledge-based management systems**: technologies, tool, and techniques from AI are directly applied to KM and KMSs.
- **Computer-supported collaborative work (groupware)**: in many parts of the world KM has become synonymous with groupware. Sharing and collaboration have become vital to organizational KM and KMSs.
- **Library and information science**: the art of classification and knowledge organization is at the core of library science, and it will become vital as more information is gathered. This science will most certainly contribute to tools for thesaurus and vocabulary management.
- **Technical writing**: technical writing, also called technical communications, is directly relevant to the effective representation and transfer of knowledge.
- **Document management**: the managing of electronic images, document management has made content accessible and reusable, becoming an essential piece in KMSs and KM activities.
- **Decision support systems**: they have brought together several disciplines, which include cognitive science, management science, computer science, operations research, and systems engineering. All of them will assist the knowledge worker or the knowledge learner in the performance of their tasks. This primarily focuses in aiding managers of organizations with their decision-making process.
- **Semantic networks**: they are knowledge representation schemes that involve nodes and links between nodes. The nodes represent objects or concepts and the links represent relations between nodes. This discipline is now in use in mainstream professional applications,
including medicine, to represent domain knowledge in an explicit way that can be shared. Thus is one of several ways that a knowledge engineer can represent knowledge.

- Relational and object databases: relational and object databases primarily contain structured and unstructured data. However, through data-mining techniques we have only begun to extract the explicit knowledge contained in these resources.

- Simulation: referred to as a component technology of KM (computer simulation) continues to contribute significantly to e-learning environments. E-learning is another key ingredient of the KMS.

- Organizational science: deals with the managing of organizations, understanding how people work and collaborate. Organizations contain many dispersed areas of knowledge where a KM policy and KMSs are essential. This discipline has led to many of the aspects involved in communities of practice and the development of communicators of practice within a KMS.

- Economics: specifically knowledge economics, which is the study of the role of knowledge in creating value because it will associate it with the valuation of the enterprise.

Interoperability can be achieved by following certain principles of development in all fields of activity. For instance, at the basics, we can start with underlying the importance of defining terms. Definitions within a community of practice can have multiple benefits, as definitions reduce differences in semantics and make knowledge accessible [6]. This helps us reach the first level of interaction and understanding between the members of a team. At the higher level, technology acts as an enabler of the collaborative practices both in learning and work environments, as computers are increasingly used to support collaborative knowledge intensive processes, where interaction takes place between team members that work together towards a common goal.

Large and complex information systems need to interoperate, in order to achieve their full potential in learning and business organisations. Interoperability brings numerous opportunities for development, driven by the possibility of interconnecting the learning environments within educational institutions with the fields of practice within business organisations. This integrated learning environment is facilitated by the service oriented architecture (SOA), which has been widely adopted to solve the interoperability of the involving heterogeneous distributed systems [7].

### Service Oriented Architecture

Service Oriented Architecture (SOA) [8] plays a key role in the integration of heterogeneous systems by the means of services that represent different system functionality independent from the underlying platforms or programming languages. SOA contributes in relaxing the complexity, leveraging the usability, and improving the agility of services. On the other hand, new services may need to be adopted by the SOA community. Service is a program that interacts with users or other programs via message exchanges. An (SOA) consists of the following concepts: application frontend, service, service repository, and service bus; each summarized as follows. Application frontends use the business processes and services within the system. A service consists of implementation, service contract, functionality and constraint specification, and service interface. A service repository stores service contracts. A service bus connects frontends to the services. A service-oriented architecture is a style of design that guides all aspects of creating and using business services throughout their lifecycle (from conception to retirement). An SOA is also a way to define and provide an IT infrastructure to allow different applications to exchange data and participate in business processes, regardless of the operating systems or programming languages underlying those applications.

### Web Services

In more technical terms, a service is a program that interacts with users or other programs via message exchanges, and is defined by the messages not by the method signatures. Web services technology is defined as a systematic and extensible framework for application-to-application interaction built on top of existing web protocols. These protocols are based on XML and include:

- Web Services Description Language (WSDL) to describe the service interfaces,
- Simple Object Access Protocol (SOAP) for communication between web services and client applications, and
- Universal Description, Discovery, and Integration (UDDI) to facilitate locating and using web services on a network.

SOAP is an XML based protocol for messaging and remote procedure call using HTTP and SMTP. It defines how typed values can be transported between SOAP representation (XML) and
application’s representation by using XML schema definition. It also defines where various parts of Remote Procedure Call (RPC) are defined, including object identity, operation name, and parameters.

WSDL has an XML format that describes web services as a collection of communication endpoints that can exchange certain messages. A complete WSDL service description has two parts: i) web service description (abstract interface), and ii) protocol-dependent details (concrete binding) that users must follow to access service at a service endpoint.

UDDI is an XML based standard that provides a unified and systematic way to find service providers through centralized registry of services.

BPEL is a language for specifying business process behaviour based on web services. These processes export and import functionality by using web service interfaces.

Web services are widely adopted as standard technology for implementation of service oriented architecture (SOA).

2.3 Mobile Technologies: a practice for unity in diversity

Collaboration is a key element of today’s world, where distributed organizations require people to collaborate across distances. Without doubt, internet services provide support for collaboration, yet these services are still limited and they need to be adapted to changing practice [5].

Even as internet services are booming, it is important to keep in mind that a knowledge process requires intense exchange of information between the team members, and in order to reach the desired outcome members need more than a simple exchange of messages or documents. They need to be able to operate interactively in a shared context.

As technology expands, and the mobile life is springing rapidly, the context of sharing is beginning to take a multi-dimensional shape and gain weight. To assimilate this rapid development, the authors have included an overview of the evolution of the mobile world in recent years. Thus, if second-generation systems like Global System for Mobile Communications (GSM) were originally designed for efficient delivery of voice services, the Universal Mobile Telecommunication Services (UMTS) networks were designed from the beginning for flexible delivery of any type of service, where each new service does not require particular network optimisation. In addition to the flexibility, the Wideband Code Division Multiple Access (WCDMA)/High-Speed Packet Access (HSPA) radio solution brings advanced capabilities that enable new services. According to Nokia experts [9], the UMTS services, for example, are divided into person-to-person services, content-to-person services and business connectivity.

Person-to-person refers to a peer-to-peer or intermediate server-based connection between two persons or a group of persons. These services refer mainly to:
- person-to-person packet switched services, such as wideband AMR, for better voice quality, video telephony, multimedia architecture for circuit-switched connections, video codec, messaging, audio messaging, instant messaging, mobile email, video sharing, push-to talk over cellular, voice over IP, or multiplayer games. Other possible data services are, for example:
  ⇒ Video conferencing. Point-to-point or multipoint session between mobiles or with office video conferencing systems.
  ⇒ Video streaming. Mobile TV with multiple live TV channels, video-on-demand for content such as news and movies.
  ⇒ Multimedia multiparty gaming. Playing in real time with other remote users.

To acknowledge the impact of these mobile technologies on our learning and working activities, let’s take a setting that allows live video or video clips in real time during a normal voice call, enabling the users to enrich their communication even more. Users can add and remove the video element as they want, sharing live camera views or video clips from the device. Below there is the flow of a video sharing session between Malcolm and Keith. Keith and Malcolm are in a standard CS call. During the ongoing CS voice call, Keith chooses to share the live video. They both have devices capable of video sharing and are registered for the service. The registration may be in always-on mode, or alternatively in when needed mode. The latter is triggered on after the CS call is connected. Keith confirms Malcolm as a recipient, then Malcolm receives video request from Keith and accepts it. The system shows the acceptance to Keith, who activates the sending of the video stream. Malcolm’s device starts showing the same video as Keith’s device. They can discuss it via the voice call. Keith ends the video sharing when he has shown what he wanted. The voice call between Keith
and Malcolm remains active. The video sharing has both professional and private use cases: sharing vacation experiences, showing real-estate property for real-estate brokers, and explaining what the situation is when there is a need to repair equipment.

This is just one of the examples that prove the utility and the impact of the improvement of mobile technologies on our professional, educational and social life.

- Content-to-person services are characterised by the access to information or download of content. They include: browsing, audio and video streaming, and content download.
- Business connectivity refers to laptop access to internet or intranet using WCDMA as the radio modem. If at the present costs represent a problem, the downward trend allows us to estimate that in the future connectivity price will lower, increasing accessibility in the learning environment.
- Location services and applications represent one of the relatively new dimensions in UMTS. It is clear that by combining location data with the communication capabilities and extensibility of a smart phone provides novel user experiences and business opportunities. Maps and city guides, including points of interest data, can always be kept up to date, which also means increased data traffic. In fact, there is a representative list of value-added services that can be updated over the air: new maps, city guides, new voice commands, weather, traffic information, safety cameras, traffic cameras and so forth.

In a future scenario, the complexity of relationships or the knowledge processes will be better supported through internet services, as mobile technologies have already been incorporated in the practice of business organisation and educational institutions.

**III. CONCLUSION**

Collaboration is a key element in providing improved performance and quality of activities both in educational and business settings. As learning escaped the formal settings and knowledge has gained ground for development, this article explores the dimensions of building collaborative systems for knowledge acquisition from semantically heterogeneous information sources. Extended access to autonomous knowledge sources will provide better resources for building optimal blends of learning components.

The authors have included insights on knowledge acquisition, promoting knowledge interoperability as a future direction of development and presenting mobile technologies as a tool for sustaining collaborative environments that comprises not only real-time, but also enriched communication in our professional, business and social life.
BIBLIOGRAPHY


* Project Manager, Advanced Technology Systems, Romania
** Software Developer, Advanced Technology Systems, Romania