RECIIS
Electronic Journal of Communication, Information & Innovation in Health
[www.reciis.cict.fiocruz.br]
ISSN 1981-6286

Technological Advance

Ontologies and information and communication technology: expert systems, the semantic web and integrated management of electronic government purchases

DOI: 10.3395/reciis.v1i1.47en

Laura Cristina Simões Viana

Vice-Presidência de Pesquisa e Desenvolvimento Tecnológico da Fundação Oswaldo Cruz – Rio de Janeiro, Brazil

Olga Fernanda Nabuco

Centro de Pesquisas Renato Archer – CENPRA Campinas, Brasil / Laboratoire D'Analyse Et D'Architecture Des Systemes – Centre National de la Recherche Scientifique LAAS/CNRS (CNRS), Toulouse, France

Carlos José Saldanha Machado

Fundação Oswaldo Cruz, Rio de Janeiro, Brazil
csaldanha@cict.fiocruz.br

Abstract

The consolidation of information and communication technology as a technical and economic paradigm suggests transformations in how science is practised. In fact, cutting edge research requires expressive processing and network capacities for handling geographically disperse data from different disciplines, marking the birth of "e-science", as in bioinformatics. In this context, this article introduces expert systems and the Semantic Web, whose construction depends on technologies for knowledge representation, such as ontologies. By spelling out a common set of concepts, these can make possible the integration of diverse and heterogeneous sets of data, facilitating the interoperability of information systems. This integration of data is a critical stage of the development of almost all computer systems, whether they are used in research activities, in new ways of organizing collaborative work, in e-science or in government or commercial activities. The article concludes by arguing that the construction of ontologies is a technological alternative for the interoperability of information systems, which may favor a new way of organizing collaborative work for the integrated management of electronic government purchases, particularly in the health sector.

Keywords

Ontologies, specialists systems, semantic web, information technology, government purchases

Introduction

When modern biotechnology, characterized by the possibility of direct manipulation of the genetic code, emerged as a technological paradigm[1] in the 1980s, microelectronics was already taking shape as a new technical and economic paradigm[2]. Initially restricted to modern research laboratories, there was speculation about the revolutionary potential of the new biotechnology and about the market which this technology would represent for microelectronics (PEREZ, 1986). Nowadays information and communication technology (ICT) based on microelectronics is considered a generic technology (FREEMAN, 2003) since as well as demonstrating innumerable technical and economic advantages, overcoming the market barrier, it has spread throughout society and is also responsible for changes in organizational and institutional dimensions. On the other hand, the most recent research in the field of biotechnology is based on "e-science", marking the arrival of bioinformatics and confirming the vocation of biotechnology as a market for ICT[3].

The term e-science[4] was coined by John Taylor from the UK Office of Science and Technology (OST) (HEY et al., 2002) to convey a new way of organizing scientific work, characterized by global collaboration in key scientific areas and its equivalent in the infrastructure necessary for operationalizing e-science. According to HEY et al. (2002), the massive number of bytes generated on a daily basis by the scientific and technological community requires efficient access to the stored data, which is geographically dispersed, as well as expressive computer and network capacity for the management and analysis of these data within feasible timescales and costs.

To meet this demand, the concept of grid computing (also known as distributed computing) emerged (HEY et al., 2002), a computer model based on distributed processing in an infrastructure which simultaneously executes the same task in various processors. The underlying idea of grid computing is analogous to that of the electric power grid, where electric energy generators are distributed and customers have access to electric energy without worrying about the source of the energy or its location. This model therefore involves a series of smaller computers, connected to each other via local networks and the Internet, which form a virtual architecture of computers which carry out high-performance tasks when idle. This new approach to networked computing is also known as metacomputing, scalable computing, global computing, Internet computing and peer-to-peer computing (BAKER et al., 2002).

E-science is not just limited to megaprojects or fundamental science. There are currently signs that e-science is changing the way in which people collaborate, how they extract knowledge from huge amounts of data, or how they organize engineering and business projects (REDFEARN, n.d.). E-science is coming to represent an innovation in the way scientific work is organized, configuring a new infrastructure for the production of scientific knowledge (HINE, 2006). However, the development of e-science and its ramifications require the integration of sets of diverse and heterogeneous data organized in different disciplines. The sharing of defined ontologies, for reasons of simplification and common concepts and their inter-relations, can help this integration.

This article introduces the idea of expert systems (ES) and the Semantic Web (SW). These technologies are examples of ICT whose development depends on the construction and the reuse of common concepts. This introduction is followed by a brief description and review of how work in artificial intelligence (AI) has evolved, an area in which the possibilities of representing knowledge are an important question for the success of the task of giving a machine the power of inference. The second section presents a general overview of expert systems and their respective applications in order to illustrate, in the following section, a new version of the Web, the Semantic Web, and how this could be put into practice. In the penultimate section, the article describes the use of the term ontology in the context of ICT and explores some of its more recent applications. Finally, the article suggests an exercise for the elaboration of an ontology of products and services which could be used to support government purchasing activities.

Artificial Intelligence and Expert Systems

Based on the founding work by TURING, MCCARTHY and MINSKY, the term artificial intelligence (AI) was coined and the foundation was laid for the use of machines which can learn and infer, making human work and life easier (MCCARTHY et al.., 1969; MCCARTHY, 2004; MACKWORTH, 2007). The area of research into symbolic AI[5] which follows the logical tradition begun by McCarthy and Hayes understands that an "intelligent" computer program must have a general representation of the world based on which its entries can be interpreted; from a practical point of view, the problem of AI is how to model "intelligence" (BITTENCOURT, 2005). The evolution of symbolic AI can be divided into three phases (BITTENCOURT, 2005): classical, romantic and modern. In the classical phase, which lasted until the 1970s, the aim of symbolic AI was to simulate human intelligence, making use of general problem-solving methods and logic. During this phase the main limitation of AI was to underestimate the computational complexity of problems. The romantic phase between the 1970s and 1980s had a similar aim of simulating human intelligence in pre-determined situations. The methods adopted in this phase were the formalisms of knowledge representation adapted to the type of problem and the mechanisms of the procedural approach[6], with the aim of achieving greater computational efficiency. There was no awareness in the romantic phase of the quantity of knowledge necessary to deal with simple common sense problems.

The emergence of expert systems (ES) marked AI's passing to the modern phase, between the 1980s and 1990s. In this phase, the aim of symbolic AI was to forge the behavior of an expert in a specific domain in solving problems. Starting in 1980 various expert systems were developed but this initial success was followed by disenchantment, reinforcing the understanding that the intelligence in an expert system is essentially to be found in the way that the knowledge about the domain is represented and that the stage of acquiring the knowledge is the most difficult part of the development of an expert system (BITTENCOURT, 2005; FURNIVAL, 1995). The following paragraphs situate this question of knowledge representation in the context of symbolic AI, describe expert systems and list some examples and their respective areas of application.

Expert systems

The classic methods (or formalisms) of knowledge representation, which have their base in logic (or first and second order logic) are production systems, semantic networks and systems of tabular representation (BITTENCOURT, 1998). Production systems include systems based on production rules, that is pairs of expressions consisting in a condition ("if") and an action ("then"). Semantic networks, used in systems for the understanding of natural language, are a set of nodes connected by sets of arches, where the former represent objects and the latter binary relations between the objects. The tables are nodes with an internal structure; this latter model is considered the basis for ideas which resulted in programming languages focused on the object. Expert systems use production rules as a method of knowledge representation.

Considering that cognitive capacity can be divided into two parts: a base of knowledge based on declarations and an inference engine (BITTENCOURT, 1998), expert systems also have two main components, a knowledge base and an inference (or reasoning) engine (or machine). The knowledge base brings together factual data or rules and the inference engine applies the rules to infer new knowledge.

The DENDRAL and MYCIN systems are examples of classical expert systems (LUGER, 2005). The DENDRAL system is emblematic of the pioneering days of expert systems focused on the resolution of problems in specific domains. Developed at Stanford at the end of the 1960s as a set of programs, the aim of the DENDRAL system was to infer the structure of organic molecules based on chemical formulae and mass spectrometry data on the chemical linkages present in the molecules (LINDSAY et al.., 1993). The DENDRAL system adopts a heuristic search method and specialized chemical knowledge when carrying out a search, which is limited to likely situations (WALKER, 1987). Although DENDRAL is known in the chemical computing community, its use in chemistry was limited and it is no longer executed as an integrated software package (LINDSAY et al.., 1993). The experience with DENDRAL proved relevant for AI and for the planning and implementation of other expert systems (LINDSAY et al.., 1993).

The MYCIN system, also developed at Stanford in the mid-1970s, launched the contemporary expert system methodology by solving reasoning problems with uncertain and incomplete information (LUGER, 2005). This system provides logical explications for its reasoning, uses a control structure adapted to the specific domain of the problem and identifies reliable criteria for evaluating its own performance (SHORTLIFFE, 1984; 1975). The MYCIN system uses specialized medical knowledge for the treatment of patients with meningitis and bacteremia (SHORTLIFFE, 1984). Together with MYCIN, the PIP, INTERNIST-1 and CASNET systems, described briefly below, are considered inaugural milestones in AI research applied to medicine (SHORTLIFFE, 1986):

· PIP System (Present Illness Program) – brings together data to generate hypotheses about illness processes in patients with kidney disease

· INTERNIST-1 System – used in the diagnosis of complex problems in internal medicine (SHORTLIFFE, 1986)

· CASNET System – an advisory system in ophthalmology used in the evaluation and management of patients with glaucoma.

At the moment, applications of AI in the practice of medicine[7] include the prescription of medication, clinical laboratories, educational contexts, clinical monitoring and areas which require the handling of a great deal of data, such as intensive care units (COINERA, 2003).

The PROSPECTOR, DIPMETER ADVISOR and XCON systems, described below, are also considered classic expert systems (LUGER, 2005):

· PROSPECTOR System – designed to locate mineral deposits such as gold and molybdenum (WALKER, 1986)

· DIPMETER ADVISOR System – used for the analysis of results gathered during oil exploration

· XCON System – used in the configuration of VAX computers (Digital Equipment Corporation).

The use of expert systems can contribute to improvements in productivity in commercial, scientific, technological and military activities. A range of commercial expert system packages are currently available on the market, with user-friendly interfaces and various applications (ENGELMORE, 1993):

· Diagnosis and identification of problems in devices and systems in various areas, such as medical and engineering systems.

· Planning and development of schedules, such as flight schedules, personnel deployment and boarding schedules and disembarkation terminal schedules for airlines, programming of manufacturing job-shops and planning of the manufacturing process.

· Configuration of manufactured objects through partial modeling such as the construction of modular houses and other activities which involve complex engineering and manufacturing projects.

· Decision-making in the financial sector, such as advisory programs used to support credit analysis, risk analysis and determination of insurance premiums, and foreign trade.

· Editing of knowledge, such as advice on spelling and grammar and tax advice, in particular for personal income tax.

· Control and monitoring of processes, such as the systems used in steel mills and oil refineries, which carry out real-time analysis of data from the equipment and other devices in order to detect anomalies, foresee trends and control the correction of errors and optimization.

· Planning and manufacturing, such as support systems for industrial design and process design both on a conceptual level and in the configuration of the factory floor in manufacturing processes.

· Information and library services, such as the systems used in indexing, automatic generation of summaries, reference works, cataloging, online information retrieval, collection development and detection of duplicate records (FURNIVAL, 1995; MENDES, 1997).

Some examples of research about expert systems carried out in Brazil include their application in medicine (RAZZOUK et al., 2006); industrial design (MAZIERO, 2000); support for decision-making in an industrial environment (HENNEMANN, 2006); the use and management of land (GIBOSHI et al., 2006); information management (MENDES, 1997); industrial process control (SELLITO, 2002); the sanitary control of seeds (ALVES et al., 2006) and the evaluation of land for planting grains (CHAGAS et al., 2006). A large part of the commercial products available on the Brazilian market which use AI combine resources from expert systems (logical) and neural networks (intuitive), with application primarily in the financial, telecommunications and public services, such as energy (ABES, n.d.).

Before moving to the next section about the Semantic Web, we should remember that the most critical stage of construction of an expert system is the acquisition of knowledge, which consists of the accumulation, transfer and transformation of the knowledge to the machine, to form the knowledge base for the expert system (CIN/UFPE, n.d.). The main knowledge acquisition stages in the construction of an expert system are the identification of the characteristics of the problem whose resolution is the object of the expert system, the development of an ontology (common concepts) in the problem domain and the identification of inferences about the concepts which make up the ontology. More recently, the acquisition of knowledge by the expert systems began to be guided by models, which include the reuse of ontologies and inference structures (CIN, UFPE, n.d.)

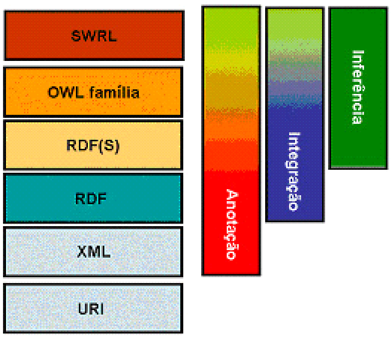
Semantic Web

Although the Web is becoming more and more like TV, it is undeniable that it remains interactive and information rich, enabling connections between an endless number of devices, such as in the intelligent house and the ability to check one's bank account from a mobile phone. However, given that the Web is no longer limited to a computer interface, there is still a way to go for the different devices and people to be able to connect to each other. Given that the Web developed as a document repository to be handled by people rather than machines, the exploration of its full potential should give priority to a language which can be understood by both people and machines. A possible solution may lie in a semantic theory which could interpret symbols, transforming data into information and providing explanations about the meanings; the logical connection between terms would establish the interoperability between systems (BERNERS-LEE et al., 2001). The Semantic Web (SW) is a new version or an expanded version of the current Web, where information is accompanied by a well-defined meaning, helping the cooperation between people and computers.

According to BERNERS-LEE et al. (2001), the functioning of the SW depends on computers having access to structured collections of information and sets of inference rules which can be used to guide the automation of the inference of automatic reasoning. As we said in section 2.0, the intelligence in a system lies in the knowledge representation. The challenge for the SW is therefore to develop a language which expresses data and reasoning rules about the data, allowing the rules of any system of knowledge representation to be exported to the Web. So what is necessary to describe information in such a way that it is equally comprehensible by people and machines? The model of layers of standards and Web technologies, shown in the diagram below and described in the following paragraphs, suggests an architecture for the SW.

The W3C recommendations[1] indicate that resources[2] pointed to by addresses should not have dubious meanings, but should rather relate data with objects in the real world. The electronic addresses for projects can be found the Uniform Resource Identifier (URI) layer, which provides the means for unambiguously identifying resources on the Web. A URI is a short string which identifies resources on the Web, such as documents, images, downloadable files, services, electronic mailboxes, and other resources. URI is a generic form of Uniform Resource Location (URL), the traditional Web address, such as in www.w3c.org.

The eXtensible Markup Language (XML) and Resource Description Framework (RDF) layers represent



Fonte: adaptado de Globe (2007)

the usual data and information formats used by the Web. The XML level, together with definitions of types and schemas, guarantees the integration of the SW definitions with other standards based on XML. RDF is the universal standard for Web data, making it possible to mix, export and share structured and semi-structured data via different applications. The Resource Description Framework describes various types of resources whereas XML schemas, for example, describe only documents; it permits interoperability between applications and allows the meaning of terms and concepts to be quickly processed by computers (Berners-Lee, 2001); it may use XML in its syntax and URI to describe entities, concepts, properties and relationships.

The RDFSchema language contains schemas which facilitate the recovery of information in the appropriate format for a specific application, making it possible for computer programs to "understand" the information. RDFSchema defines classes and properties specific to applications, where classes may be collections of resources. The Web Ontology Language (OWL) layer uses description logics (first-order logic) and aims to extend the meaning of applications which use the Web to include programs which interpret this information automatically. The objective is to introduce the interoperability of machines, automatic programs, agents[3], etc. with Web systems. The OWL language makes it possible to make inferences about the content which it represents. Semantic Web Rule Language (SWRL) is a language for writing rules in first-order logic. Used in conjunction with OWL, it forms a powerful pair of knowledge representation languages for the Web. SWRL and OWL together make it possible to store and recover categorized language. In addition, they also allow an inference engine to use this base to add new knowledge to its own base.

Annotations are made about the main knowledge base by adding a link to information which explains the information in more detail. The Annotations layer adds information to the information, making it possible to distinguish it from other information and describing it better. The Integration layer ports interoperability to applications, integrating information sources; in semantic terms, the integration brings to the surface the exclusive meaning of the information, which extrapolates that of a simple key word or a thesaurus, moving in the direction of a "yellow pages"-type catalog. Each layer increases the recognition of the value of the information. Inference is possible when the description language is rich enough to describe axioms and logical phrases, allowing a program, such as an expert system or a rule engine, to interpret the information, adding new information to it. This new information, although it is not yet written, obeys the existing rules. The Inference layer carries out reasoning based on the existing information, and creates new instances of information. This layer does not create new rules, such as the conclusions constructed based on transitivity: if Flávio Simões is the son of Luiz Simões and each person can have only one father and João Barbosa is Flávio Simões' father then, the machine concludes, Luiz Simões and João Barbosa are the same person. To summarize, the development of the Semantic Web concentrates on the definition of layers of languages, used to support the representation and use of metadata; the languages are the basic tool used to add meaning to the information necessary for the Semantic Web (GLOBE, 2007).

There are innumerable Semantic Web applications and the suggestions which follow are just a sample: metadata library (for integration in a single domain); integration of data in public safety; integration of research and development data; search tools; connections between knowledge from different domains, such as genomics, proteomics, clinical trials, rules and regulations and others; electronic government; and energy (HERMAN, 2007). In terms of Internet pages, the Mind Lab[4] at the Institute for Advanced Computer Studies at the University of Maryland combines Web technologies (HTML, XHTML, XML, PHP, CSS, etc.) with SW languages (RDF, RDFS, DAML+OIL, OWL) and other tools. The Friend of a Friend (FOAF)[5] project, which sets out to create personal Web pages which are also machine readable also incorporates SW technologies. The Biotechnological Center (BIOTEC) at the *Technische Universität Dresden* developed a search engine for the life sciences, in particular molecular biology, called GoPubMed[6], based on the PubMed service[7]. The search results are classified according to GeneOntology[8], a controlled vocabulary for describing the gene and the attributes of its products in any organism. At the same time, some commercial products already use SW technologies[9] developed by the W3C Consortium.

These initiatives are still some distance away from a model of the Web based on data and information, in other words, from the Semantic Web. This distance exists because there is still no mediation based on agents on a large scale, on which the consolidation of Web standards and technologies which describe meanings depends (SHADBOLT et al., 2006). On the other hand, the growing use of ontologies in e-science for the integration of sets of different and heterogeneous data, originating in different areas, may accelerate the development of the SW (SHADBOLT et al., 2006). Other experts, such as LÉVY (2006), argue, however, that the SW does not solve the problem of semantic interoperability[10] because the annotation of concepts in natural language is arbitrary and also because the numerous ontologies are incompatible.

The meaning of Ontology

Two inter-related questions are present in many of the discussions about artificial intelligence: knowledge representation and interoperability. The former, already covered in section 2.0, deals with the representation of the world (or part of it) in such a way that machines can process it. This question forms a specific area of study - knowledge engineering – which resulted from the development of expert systems (WELTY, 2003). The other more recent question concerns interoperability between information systems, which still has a long way to go before communication between systems, as well as between them and their users, will be possible. In both questions, ontologies play a fundamental role.

There are many constructions called ontologies which differ in the design of the modeling, in the representation used for the modeling and the philosophical point of view (BODENRIDER et al., 2006). The term ontology, which originates in philosophy, was incorporated by computer science at the beginning of the 1980s, when John McCarthy used it for the first time (WELTY, 2003). Since then, many definitions for ontology have emerged without news, so far, of a consensus about the term in computer science.

According to WELTY (2003), ontology in computer science deals with the meaning and the existence of objects and concepts; for the author, ontology defines the objects and concepts which exist within the domain of a system, their relationship to each other and the clearest possible meaning of these objects and concepts. For GRUBER (1983), the important thing is the purpose of the ontology, and in the context of knowledge sharing and reuse, ontology can be defined as an explicit specification of a concept. In other words, ontology is a description of concepts and the relationships which may exist between them. In this context, the specification of the ontology creates ontological commitments, which stand for an agreement about the adoption of a consistent vocabulary, even if incomplete, in relation to the theory specified by the ontology. According to GRUBER (1983), conceptualization is the basis for a body of knowledge which is formally represented. For STUDER et al. (1998) conceptualization refers to an abstract model of a phenomenon, which identifies the concepts which are relevant to this phenomenon; the explicit specification indicates that the types of concepts used and the restrictions related to the use of the respective concepts are explicitly defined.

According to SHADBOLT et al. (2006), over the past five years, Semantic Web projects have pointed to the need to develop, manage and validate ontologies, independent of their domain. The reuse of ontologies and data should lead users to the reuse and discovery of information, movements which can raise big challenges for research around the Semantic Web. For example, it might be necessary to look for help from various areas, such as social network analysis and epidemiology, to understand how information and concepts spread through the Web and how to guarantee the origin and the reliability of information and concepts (SHADBOLT, 2006). The developments, methodologies, challenges and techniques under discussion about the Semantic Web will contribute to the construction of a new Web and, according to SHADBOLT (2006) and BERNERS-LEE et al. (2006), to a new discipline: Web Science. Web Science will seek to develop, implement and understand distributed information systems, systems of people and systems of machines operating on a global scale (SHADBOLT, 2006).

Practical applications of ontologies

As was mentioned in the introduction to this article, research, development and innovation projects in modern biotechnology are multidisciplinary and require the integration of sets of diverse and heterogeneous data, becoming one of the main markets for information and communication technology and marking the beginning of bioinformatics. This integration may be favored by the adoption of common ontologies. In recent years, the development of ontologies for specific use in particular domains has become a reality, allowing specialists from one domain to share and annotate information in their respective areas of expertise (NOY et al., 2001). Ontologies are being adopted not just by scientific communities but also in commercial[11] and governmental activities, as a way of sharing, reusing and processing knowledge about a domain.

For example, in the area of medicine[12] there have been two relevant developments: the Systematized Nomenclature of Medicine (SNOMED)[13], which is a structured and standardized vocabulary, and the Unified Medical Language System (UMLS)[14], a semantic network. In the biomedical area, the Open Biomedical Ontologies (OBO)[15] page brings together structured vocabularies for shared use among different domains in the biological and medical areas. The OBO contains generic ontologies[16] applied to all the organisms and ontologies of a more restricted scope, to be used in specific taxonomic groups. The Gene Ontology (GO)[17] project offers a controlled vocabulary for describing genes and the attributes of gene products in any organisms. Other disciplines are also seeking to integrate data and information, following the example of the Plant Ontology™ Consortium[18] which develops, accompanies and shares ontologies about plant structures and their stages of growth and development, and the Marine Metadata Interoperability[19] project which promotes the exchange, integration and use of data in the marine sciences. The development of the Lattes Platform Markup Language (LMPL in the Portuguese acronym)[20] is a

Brazilian initiative for the integration of information systems in science and technology through the development of a common ontology in this domain (Pacheco et al., 2001).

In the governmental area, the European Commission's 2003/98/EC Directive established general conditions for the reuse of documents produced by member countries' public sectors and access to these documents. This Directive aims to facilitate the wide circulation of information about the products and services contained in the documents, including in the business world, thereby promoting competition (EU, 2003). In the United Kingdom the establishment of policies relating to standards, access and encouragement of the reuse of public sector information are the responsibility of the Office of Public Sector Information (OPSI)[21]. Also in the United Kingdom, the GovTalk[22] page is a shared forum for the development of policies and standards for electronic government; the Electronic Service Delivery (ESD)[23] page makes available controlled lists and standards of associated data, used in all areas of the public sector; the Integrated Public Sector Vocabulary (IPSV)[24] presents a encoding scheme, used in the marking of information, aiming to guarantee a transparent flow of information between public sector bodies and offer citizens and businesspeople better access to public services.

General-use ontologies can also be found, such The United Nations Standard Products and Services Code® (UNSPSC)[25] taxonomy, which brings together an open classification for products and services in general. The UNSPSC code is an open standard business tool, and its specifications are therefore available. The aim of this code is to support the purchasing activities of businesses and institutions, consolidating values and relating them to products and suppliers. The UNSPSC code is compatible with other systems and permits aggregate and disaggregate analyses at different stages of the purchasing process, including electronic purchases. The segments included in the UNSPSC are: raw materials, industrial equipment, final products, services and components.

Exercise: elaboration of an ontology for managing electronic government purchases

Before coming to the end of this article we would like to suggest an exercise based on the elaboration of an ontology which can be used by Brazilian public sector bodies as a tool to support electronic government purchasing activities, which represent 57% of the total purchases made by the federal government (BRASIL, 2007). It is estimated that 2 billion Brazilian *reais* (R$), or 22% of this total, relate to the health sector[26].

According to article 37, clause XXI of Brazil's Federal Constitution, the works, services, purchases and sales in the area of the Brazilian public administration must be contracted through a public tender process, which guarantees equality of opportunity to all competitors. The basic regulatory framework for government purchasing activities includes:

· Law nº 8.666 of 21 June 1993, which regulated article 37, clause XXI of the Federal Constitution, and put in place standards for tenders and contracts in the public administration (BRASIL, 1993)[27].

· Decree nº 3.555 of 8 August 2000, which regulated the *Pregão* bidding method for the federal administration (BRASIL, 2000).

· Law nº 10.520 of 17 June 2002, which established in the context of the Union, States, Federal District and Municipalities, also in the terms of article 37, clause XXI of the Federal Constitution, the tender method called *Pregão*, for the acquisition of ordinary goods and services (BRASIL, 2002).

· Decree nº 5.450 of 31 May 2005, which made the *Pregão* compulsory in public contracts in the federal government, giving preference to the electronic version (BRASIL, 2005).

The *Pregão* tender method is used for the acquisition of ordinary goods and services, covering those whose performance and quality standards may be objectively defined through specifications which are well-known in the market. This group includes around 50,000 types of products and 2500 service modalities (BRASIL, 2007). The Electronic *Pregão* is carried out entirely on the Internet, through the federal government purchasing portal, *Comprasnet*[28]. According to federal government data (BRASIL, 2007), the country saved R$ 1.8 billion with the Electronic *Pregão* in 2006. In the same year, the Brazilian government acquired R$ 11.1 billion worth of ordinary goods and services through the electronic method. In total, the federal government contracted R$ 19.6 billion of ordinary goods and services in 2006. As well as reducing costs for the administration, the Electronic *Pregão* can speed up and simplify the contractual process and increase the security, transparency and democratization of government purchases, since these take place via the Internet.

A study carried out by the World Bank (BRASIL, 2006) indicates that *Comprasnet*, the Brazilian system for electronic government purchases, is efficient in terms of transparency in the publication of tenders and in their respective results, and in the use of competitive tender methods. This same study recommends improvements in the integration of *Comprasnet* with the systems for managing contracts and payments to suppliers, widening the reach of the system over purchasing logistics. In addition, it is known that the names given to products and services by end users and managers of the purchasing process vary. This practice slows things down and makes the integration of information systems between and within government bodies difficult. In particular, although the *Comprasnet* system has a Catalog of Materials, this has not yet been completely incorporated by the government bodies and is not known to a large part of the end users and managers of purchasing activities. The *Comprasnet* Catalog of Materials uses the Federal Supply Classification[29], created and developed by the United States Department of Defense, with the aim of establishing and maintaining a uniform system of identification, codification and cataloging for all the bodies which make up its structure.

The development of an ontology of products and services for purchasing activities in general, and those of the health sector in particular, could facilitate the practice of electronic purchasing and the integration between the teams responsible for the planning, execution and management of purchasing within and between government bodies.

Summary and Conclusion

The consolidation of information and communication technology (ICT) was presented throughout this article as a generic technology (FREEMAN, 2003) which has been responsible for expressive transformations in the technical, economic, organizational and institutional dimensions of contemporary societies. The shaping of e-science in recent years has confirmed the vocation of advanced scientific activities as a market for ICT and shows ICT's potential for transformation. E-science requires massive processing and network capacity to handle and manage the huge amount of diverse and heterogeneous data which is generated by scientific work. This comes to be organized in a new collaborative infrastructure for the production of scientific knowledge, with reflections in other technical and commercial activities. The evolution of this new way of organizing collaborative work requires more powerful machines and networks in arrangements such as grid computing, but also the integration of data and systems. The development of ontologies, which have application in expert systems and the Semantic Web, is a technological alternative for the interoperability of information systems and may work in favor of this new way of organizing collaborative work in support of an integrated management of electronic government purchases, particularly in the health sector.

Notes

1 The concept of "technological paradigm" (DOSI, 1982) is equivalent to that of "natural trajectories" (Nelson et al., 1977) and corresponds to the logic which guides the evolution of a specific technology.

2 Although the concept of "technical and economic paradigm" is similar to those of "technological paradigm" and "technological trajectory", FREEMAN (1990) argues that a technical and economic paradigm includes orienting principles which evolve into the common sense of engineers and managers during the development of a new cycle of economic growth.

3 The global bioinformatics market was estimated at US$ 38 billion in 2006 (FREEMAN, 2003).

4 The term cyberinfrastructure (NSF, 2007) is the North American equivalent of the term e-science which has been adopted in the United Kingdom.

5 The three main areas of research in AI are: symbolic AI, connectionist AI and evolutionary computing. Connectionist AI works on the modeling of human intelligence, simulating the functioning of neurons and their interconnections. Evolutionary computing is based on evolutionary mechanisms found in nature (BITTENCOURT, 2005).

6 The procedural approach handles and reasons about specific problems, in simple worlds, in order to avoid the problems of combinatory explosion which are typical of general methods (BITTENCOURT, 1998).

7 The http://www.openclinical.org/applications.html website provides information about knowledge management systems, systems to support decision-making and clinical applications designed for health professionals.

8 The World Wide Web Consortium (W3C), http://www.w3c.org, develops interoperable technologies (specifications, guidelines, software, and tools) to lead the Web to its full potential. W3C is a forum for information, commerce, communication, and collective understanding.

9 Resources include any entity, such as Web pages, parts of a Web page, devices, people and others.

10 Agent(s): part of a computer program executed without direct control or human supervision to reach targets established by the user. Agents may collect, filter and process information found on the Web, including with the help of other agents (BERNERS-LEE, 2001).

11 Mind Lab: http://www.mindswap.org

12 Friend of a Friend: http://www.foaf-project.org/index.html

13 GoPubMed: http://www.gopubmed.org/

14 PubMed: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=PubMed&itool=toolbar. PubMed is a service of the US National Library of Medicine which includes 16 million citations in MEDLINE and other life sciences periodicals.

15 GeneOntology: http://www.geneontology.org/

16 A list of the products can be found at: http://esw.w3.org/topic/CommercialProducts#head-5ef4570c3401e6fbb8c719b398fc1692b8535d74

17 According to Pierre Lévy (2006), semantic interoperability is equivalent to the development of digitally-based collective intelligence.

18 We found ontologies of the taxonomy type in Web page categories, such as those on the Yahoo! site and those for products on sale and their characteristics, as on the Amazon webiste (Noy et al., 2001).

19 Bodenreider et al. (2006) present a review of the current status of ontologies in medicine.

20 SNOMED: http://www.snomed.org/snomedct/

index.html

21 UMLS: http://www.nlm.nih.gov/pubs/factsheets/umls.html

22 OBO: http://obo.sourceforge.net/

23 OBO – specific generic ontologies: http://obo.sourceforge.net/cgi-bin/table.cgi

24 GO: http://obo.sourceforge.net/cgi-bin/table.cgi

25 Plant OntologyTM Consortium: http://www.plantontology.org/

26 Marine Metadata Interoperability project: http://marinemetadata.org/

27 Lattes Platform: http://lattes.cnpq.br/index.htm

28 OPSI – http://www.opsi.gov.uk/about/index.htm

29 GovTalk: http://www.govtalk.gov.uk/howitworks/howitworks.asp

30 ESD: http://www.esd.org.uk

31 IPSV: http://www.esd.org.uk/standards/ipsv_internalvocabulary/

32 UNSPSC: http://www.unspsc.org/

33 This estimate is based on the total value of tenders in the *Pregão* bidding method in 2006 (R$ 8,833,380,000) and the value corresponding to the Ministry of Health in the same period (R$1,945,389,000), available from http://www.comprasnet.gov.br/publicacoes/boletins/12-2006.pdf.

34 A proposed law which would alter the provisions of this law is currently being processed by the National Congress. For example, it is being proposed that any tender method established in Law no 8.666 can be carried out and processed through an electronic system linked to the Internet (BRASIL, 2007a).

35 Brazilian electronic portal for government purchasing: http://www.comprasnet.gov.br.

36 http://www.dlis.dla.mil/default.asp

Bibliographic references

ALVES, M.C. et al. Desenvolvimento e validação de um sistema especialista para identificar fungos na análise sanitária de sementes. **Revista Brasileira de Sementes**, v.28, n.1, p.176-186 (2006). Available at: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0101-31222006000100025&lng=pt&nrm=iso>. Accessed: 7 Apr. 2007.

ASSOCIAÇÃO BRASILEIRA DAS EMPRESAS DE SOFTWARE – ABES (s.d.) *Entrevista com Daniel Boacnin: criatividade para crescer*. s.l., s.p. Available at: <http://www.abes.org.br/default.aspx>. Accessed: 8 Apr. 2007.

BAKER, M.; BUYYA, R.; LAFORENZA, D. (2002) Grids and Grid technologies for wide-area distributed computing. **Software – Practice and Experience**. v.32, n.15, p.1411-1506, Dec. 2002. Available at: <http://www.gridbus.org/papers/gridtech.pdf>. Accessed: 23 Mar. 2007.

BERNERS-LEE, T. et al. Creating a science of the Web. ***Science***, v.313, n.5788, p.769-771 (2006). Available at: <http://www.sciencemag.org/cgi/content/full/313/5788/769?maxtoshow=&HITS=10&hi...>. Accessed: 18 Apr. 2007.

BERNERS-LEE, T. *Web for real people*. s.l.:W3C. s.p. (2005). Available at: <http://www.w3.org/2005/Talks/0511-keynote-tbl/>. Accessed: 7 Apr. 2007.

BERNERS-LEE, T.; HENDLER, J.; LASSILA, O. The semantic web: a new form of web content that is meaningful to computers will unleash a revolution of new possibilities. **Scientific American**, p.34-43, May 2001.

BITTENCOURT, G. Representação de conhecimento: da metafísica aos programas. In: SBC. (Org.). **Cursos JAI'98.** Belo Horizonte: UFMG. (1998) p.283-333. Available at: <http://www.das.ufsc.br/~gb/pg-ia/rdec.pdf>. Accessed: 9 Mar. 2007.

BITTENCOURT, G. Computação evolutiva. In: *CICLO DE CURSOS E PALESTRAS DA ENGENHARIA ELÉTRICA, 11*. CIPEEL 11., 2005, Santa Catarina. UFSC: Florianópolis. p.23. Available at: <http://www.das.ufsc.br/~gb/CIPEEL/cevol.pdf>. Accessed: 9 Mar. 2007.

BONDENREIDER, O.; STEVES, R. Bio-ontologies: current trends and future directions. **Briefings in Bioinformatics**, v.7, n.3, p.256-274, 2006.

BRASIL. Lei n. 8.666, de 21 de junho de 1993. Regulamenta o art. 37, inciso XXI, da Constituição Federal, institui normas para licitações e contratos da Administração Pública e dá outras providências. Available at: <http://www.planalto.gov.br/ccivil/Leis/L8666cons.htm>. Accessed: 12 Jan. 2007.

BRASIL. Decreto n. 3.555, de 08 de agosto de 2000. Aprova o regulamento para modalidade de licitação denominada pregão, para aquisição de bens e serviços comuns. Available at: <https://www.planalto.gov.br/ccivil_03/decreto/d3555.htm>. Accessed: 12 Jan. 2007.

BRASIL. Lei n.10.520, de 17 de junho de 2002. Institui, no âmbito da união, estados, distrito federal e municípios, nos termos do art. 37, inciso xxi, da constituição federal, modalidade de licitação denominada pregão, para aquisição de bens e serviços comuns, e dá outras providências. Available at: <http://www.planalto.gov.br/CCIVIL/leis/2002/L10520.htm>. Accessed: 12 Jan. 2007.

BRASIL. Decreto n. 5.450, de 31 de maio de 2005. Regulamenta o pregão, na forma eletrônica, para aquisição de bens e serviços comuns, e dá outras providências. Available at: <https://www.planalto.gov.br/ccivil_03/_Ato2004-2006/2005/Decreto/D5450.htm>. Accessed: 12 Jan. 2007.

BRASIL. *Notícias Comprasnet*. 29/11/2006 - Estudo do banco mundial atesta eficiência do sistema comprasnet. 29 Nov. 2006. Available at: <http://www.comprasnet.gov.br/noticias/noticias1.asp?id_noticia=216>. Accessed: 3 Apr. 2007.

BRASIL. *Notícias Comprasnet*. Pregão eletrônico gera economia de R$ 1,8 bilhão em 2006. 16 Jan. 2007. Available at: <http://www.comprasnet.gov.br/noticias/noticias1.asp?id_noticia=219>. Accessed: 3 Apr. 2007.

BRASIL. *Notícias Comprasnet*. Pac contempla alterações na lei geral de licitações. 23 Jan. 2007a. Available at: <http://www.comprasnet.gov.br/noticias/noticias1.asp?id_noticia=219>. Accessed: 3 Apr. 2007.

CENTRO DE INFORMÁTICA - CIN/UFPE (s.d.) *Engenharia de Conhecimento e Sistemas Especialistas*. Recife. Available at: <www.di.ufpe.br/~compint/aulas-IAS/se.ppt>. Accessed: 11 Mar. 2007.

CHAGAS, C. S. et al. Aplicação de um sistema automatizado (ALES - Automated Land Evaluation System) na avaliação das terras das microrregiões de Chapecó e Xanxerê, Oeste Catarinense, para o cultivo de grãos. **Revista Brasileira de Ciência do Solo**, v.30, n.3, p.509-522, 2006. Available at: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-06832006000300>. Accessed: 5 Apr. 2007.

COIERA, E. Clinical decision support systems. In: **Guide to health informatics** 2 Ed. Sydney: Arnold. (2003) Cap. 25, s.p. Available at: <http://www.coiera.com/aimd.htm>. Accessed: 4 Apr. 2007.

DOSI, G. Technological paradigms and technological trajectories. a suggested interpretation of the determinants and directions of technical change, **Research Policy**, v.11, n.3, p.147-208, Jun. 1982.

ENGELMORE, R.S.; FEIGENBAUM, E. Introduction. In: ENGELMORE, R.S. (Ed.) **Knowledge-based systems in Japan. s.l.**: WTEC Hyper-Librarian, 1993. Available at: <http://www.wtec.org/loyola/kb/toc.htm>. Accessed: 30 Mar. 2007.

EUROPEAN COMMISSION. Directive 2003/98/EC, de 17 novembro de 2003. *Official Journal of the European Union*, Bruxelas: EC. p.7. Available at: <http://europa.eu.int/eur-lex/pri/en/oj/dat/2003/l_345/l_34520031231en00900096.pdf>. Accessed: 12 Apr. 2007.

FREEMAN, C. *Policies for developing new technologies*. Sussex: SPRU. 18p. SEWPS, n.98, 2003. Available at: <http://www.sussex.ac.uk/Units/spru/publications/imprint/sewps/sewp98/sewp98.pdf>. Accessed: 6 Apr. 2007.

FREEMAN, C. Prometheus unbound. In: **The economics of innovation.** Aldershot: Edward Elgar Pub. Ldt. 1990. Parte V, Cap 26, p.494-507.

FURNIVAL, A.C. Delineando as limitações: sistemas especialistas e conhecimento tácito. **Ciência da Informação,** v.24, n.2, s.p. 1995. Available at: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-73862000000400003&lng=en&nrm=iso>. Accessed: 9 Mar. 2007.

GIBOSHI, M.L.; RODRIGUES, L.H.A.; LOMBARDI NETO, F. Sistema de suporte à decisão para recomendação de uso e manejo da terra. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v.10, n.4, p.861-866, 2006. Available at: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1415-662006000400012&lng=pt&nrm=iso>. Accessed: 7 Apr. 2007.

GLOBE, C. Sense and Sensibility: Semantic Infrastructure for Bioinformatics. In: INITIATIVE FOR INNOVATIVE COMPUTING. Harvard. 31 Jan. 2007. Available at: <http://www.cs.man.ac.uk/~carole/talks/index.html>. Accessed: 7 Abr. 2007.

GROSOF, B.N. et al. Description logic programs: Combining logic programs with description logic. In: PROC. OF THE TWELFTH INTERNATIONAL CONFERENCE ON WORLD WIDE WEB, Budapest: ACM, 20-24 May. p.48-57, 2003.

GRUBER. T.R. A translation approach to portable ontologies. **Knowledge Acquisition**, v.5, n.2, p.199-220, 1993.

HENNEMANN, F.A. et al. Um sistema híbrido de apoio à decisão formado por redes de Petri, simulação e sistema especialista. **Revista Controle & Automação**, v.17, n.1, p.10-23, 2006. Available at: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0103-17592006000100002&lng=en&nrm=iso>. Accessed: 7 Apr. 2007.

HERMAN, I. Introduction to the Semantic Web: tutorial on Semantic Web technologies. In: INTERNATIONAL CONFERENCE ON SEMANTIC WEB & DIGITAL LIBRARIES. Bangalore. 21 Fev. 2007. Available at: <http://www.w3.org/2007/Talks/0221-Bangalore-IH/>. Accessed: 9 Apr. 2007.

HEY, T.; TREFETHEN, A.E. The UK e-science core programme and the Grid. **Future Generation Computing Systems**, v.18, p.1017-34, 2002. Available at: <http://www.rcuk.ac.uk/cmsweb/downloads/rcuk/research/esci/coreproggrid.pdf>. Accessed: 29 Mar. 2007.

HINE, C.M. Preface. In: **New infrastructures for knowledge production; understanding e-science**. London: INFOSCL. 2006, p.vii-xvii.

LÉVY, P. **IEML: computational semantics in the service of collective intelligence**. Otawa: CRC/FRSC. p.9, 2006 (translated by Michele Healy). Available at: <http://www.ieml.org/IMG/pdf/vision-ieml-Initiative-en.pdf>. Accessed: 8 Apr. 2007.

LINDSAY, R.K. DENDRAL: a case study of the first expert system for scientific hypothesis formation. **Artificial Intelligence** v.61, p.209-261, 1993. Available at: <http://profiles.nlm.nih.gov/BB/A/B/O/M/_/bbabom.pdf>. Accessed: 8 Apr. 2007.

LUGER, G.F. Artificial intelligence: its roots and scope. In: ***Artificial intelligence*: structures and strategies for complex problem solving**. 5 ed. s.l.: Addison-Wesley. 2005. p.1-34. Available at: <http://www.cs.unm.edu/~luger/ai-final/chapter1.html>. Accessed: 30 Mar. 2007.

MACKWORTH, A.K. AI Overview. AAAI, 2007. Available at: <http://www.aaai.org/AITopics/html/overview.html>. Accessed: 24 Mar. 2007.

MAZIERO, N. L. et al. A feature-based object-oriented expert system to model and support product design. **Journal of the Brazilian Society of Mechanical Sciences**, v.22, n.4. s.p., 2000. Available at: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-73862000000400003&lng=en&nrm=iso>. Accessed: 7 Abr. 2007.

MCCARTHY, J.; HAYES, P.J. (1969) Some philosophical problems from the standpoint of artificial intelligence. In: MICHIE, D.; MELTZER, B. (Eds.) **Machine intelligence**. Edinburgh: Edinburgh University Press. p.463-502. Available at: <http://www-formal.stanford.edu/jmc/mcchay69/mcchay69.html>. Accessed: 24 Mar. 2007.

MCCARTHY, John. (2004) **What is artificial intelligence?** Standford: Computer Science Department Stanford University. Available at: <http://www-formal.stanford.edu/jmc/whatisai.html>. Accessed: 24 Mar. 2007.

MENDES, Raquel Dias. (1997) Inteligência artificial: sistemas especialistas no gerenciamento da informação. **Ciência da Informação**, v.26, n.1, p.39-45, Jan.-Abr. 1997. Available at: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-19651997000100006&lng=en&nrm=iso>. Accessed: 7 Apr. 2007.

NELSON, R.R.; WINTER, S.G. (1977) In search of useful theory of innovation, **Research Policy**, v.6, n.1, p.36-76, Jan. 1977.

NOY, N.F.; McGUINNESS, D.L. (2001) **Ontology Development 101: A Guide to Creating Your First Ontology**. S. Stanford University. p.25. Available at: <http://www.ksl.stanford.edu/people/dlm/papers/ontology101/ontology101-noy-mcguinness.html>. Accessed: 9 Mar. 2007.

NSF (2007) **Cyberinfrastructure vision for 21st century discovery**. Arlington. 64 p. Available at: <http://www.nsf.gov/pubs/2007/nsf0728/nsf0728.pdf>. Accessed: 8 Apr. 2007.

PACHECO, Roberto Carlos dos Santos; KERN, Vinícius Medina (2001) Transparência e gestão do conhecimento por meio de um banco de teses e dissertações: a experiência do PPGEP/UFSC. **Ciência da Informação**, v.30, n.3, p.64-72, Sep.-Dec. 2001. Available at: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-19652001000300009&lng=pt&nrm=iso>. Accessed: 6 Mar. 2007.

PEREZ, C. (1986) Las nuevas tecnologías; una visión de conjunto. In: OMINAMI, C. (Ed). **La tercera revolución industria;** impactos internacionales del actual viraje tecnológico. Buenos Aires: Grupo Editor Latinoamericano. cap.2, p.43-89

RAZZOUK, D. et al. (2006) Decision support system for the diagnosis of schizophrenia disorders. **Brazilian Journal of Medical and Biological Research**, v.39, n.1, p.119-128. Available at: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-879X2006000100014&lng=en&nrm=iso>. Accessed: 7 Apr. 2007.

REDFEARN, J. (s.d.) **Highlights from de UK e-science programme**. The Research Councils UK: s.l. 4p. Available at: <http://www.rcuk.ac.uk/cmsweb/downloads/rcuk/research/esci/escihighlights.pdf>. Accessed: 23 Mar. 2007.

STUDER, R.; BENJAMINS, R.; FENSEL, D. Knowledge engineering principles and methods. **Data and Knowledge Engineering** (DKE), v.25, n.1/2, p.161-197, 1998.

WELTY, C. Ontology Research. **AI Magazine**, v.24, n.3, p.1-12, 2003. Available at: <http://www.aaai.org/AITopics/assests/PDF/AIMag24-03-002.pdf>. Accessed: 9 Apr. 2007.

About the authors

Laura Viana

Laura Viana has a Doctorate on Production Engineering by the *Coordenação dos Programas de Pós-graduação de Engenharia* (Coppe) of the *Universidade Federal do Rio de Janeiro*. Currently, she is an analyst in Science and Technology at *Fundação Oswaldo Cruz*. She has professional experience in the public and private sectors, with emphasis in technological innovation and industrial organization applied to information and communication technology industries, pharmaceutical and food industries, in studies and projects about the innovation in services, research and technological development and public policies.

Olga Nabuco

Olga Nabuco has a Doctor degree in Mechanic Engineering at the *Universidade Estadual de Campinas*. She is a researcher at the *Centro de Pesquisa Renato Archer* [Renato Archer Research Center]. Currently, she is a visiting researcher at the *Laboratoire D'Analyse Et D'Architecture Des Systemes – Centre National de la Recherche Scientifique (LAAS/CNRS)*. Her area of activity is characterized, mainly, by studies on knowledge representation models, ontology, knowledge sharing, service-oriented architecture, business models and web semantics.