ProLink: A Semantics-based Social Network for Software Projects

Juan Miguel Gómez Berbis,

Ricardo Colomo Palacios*,

Ángel García Crespo,

Belén Ruiz Mezcua

Computer Science Department, Carlos III University,
Av Universidad 30, Leganés (Madrid), SPAIN
Fax: +34 91 624 94 30

E-mail: {juanmiguel.gomez, ricardo.colomo, angel.garcia, belen.ruiz}@uc3m.es

*Corresponding author

Abstract: One of the core challenges in the IT world is to harness the possibilities of reusing knowledge, expertise and lessons learned from previous software projects. In this paper, we present ProLink, a semantic-based social network, a proof-of-concept architecture and implementation of how this approach could foster expertise sharing and its potentially tremendous impact on work organisation and software development.

Keywords: Social Networks, Semantics, Competency, Software Projects, Staffing

Biographical notes: Juan Miguel Gomez is a Visiting Professor at the Computer Science Department of the Universidad Carlos III de Madrid. He holds a PhD in Computer Science from the Digital Enterprise Research Institute (DERI) at the National University of Ireland, Galway and received his MSc in Telecommunications Engineering from the Universidad Politécnica de Madrid (UPM). He was involved in several EU FP V and VI research projects and was a member of the Semantic Web Services Initiative (SWSI). His research interests include semantic web, semantic web services, business process modelling, b2b integration and, recently, bioinformatics.

Ricardo Colomo has been a Faculty Member of the Computer Science Department at Carlos III Technical University of Madrid since 2002. His research interests include Software Process Improvement, Software Project Management

and Information Systems. He received his PhD in Computer Science from the Universidad Politécnica of Madrid (2005). He also holds a MBA from the Instituto de Empresa (2002). He has been working as software engineer, project manager and software engineering consultant in several companies including Spanish IT leader INDRA.

Angel García-Crespo is the Head of the SofLab Group at the Computer Science Department in the Universidad Carlos III de Madrid and the Head of the Institute for promotion of Innovation Pedro Juan de Lastanosa. He holds a PhD in Industrial Engineering from the Universidad Politécnica de Madrid (Award from the Instituto J.A. Artigas to the best thesis) and received an Executive MBA from the Instituto de Empresa. Professor García-Crespo has led and actively contributed to large European Projects of the FP V and VI, and also in many business cooperations. He is the author of more than a hundred publications in conferences, journals and books, both Spanish and international.

Belen Ruiz Mezcua is a Lecturer at the Computer Science Department of the Universidad Carlos III de Madrid. She holds a PhD in Sciences Physics from the Telecommunications School of Polytechnic University of Madrid. She is sub-head of Institute for promotion of Innovation Pedro Juan de Lastanosa. Her research interests are focused on speech and she works in, and manages, several International Projects. She is co-author of several publications in refereed international journals and conferences.

1. Introduction

Software development is hardly a linear process. Failure rates in software projects are high and the qualified software engineers able to deal with software development processes and their shortcomings and caveats (Pressman, 2005) represent a scarce resource. The heterogeneous training, background and expertise required to the professionals who join software development teams (McConnell, 2003) pose special challenges to human resource managers, who are called to apprise complex competences and stimulate their continuous development (Curtis, Hefley & Miller, 2001).

At the same time, the ever-changing nature of Information Technology (IT) provides the ideal test bed for acquiring the needed knowledge, experience and expertise at different organisations. This can be achieved by means of sharing resources, for example, through social networks. The conceptualisation of social networks as distributed repositories of knowledge is not new, yet the availability of technology that allows tracking required competences, skills and expertise could greatly enhance their potential.
In this work, we argue that semantic technology is an effective solution to find, analyze and identify skills, experience and expertise, and, more practically, software projects sharing similar challenges and goals. We present ProLink, a semantics-based social network that fosters the potential of semantics and social networks to reuse, share and gain the aforementioned knowledge, experience and expertise.
The remainder of this paper is organized as follows. Section 2 surveys the relevant literature and presents a conceptual model for competence and competence management. Section 3 describes social networks, discusses ongoing IT-based developments and presents

the *Description of a Project* (DOAP) ontology as a useful knowledge source. Section 4 sketches how the ProLink architecture and implementation fits in the picture. Finally, Section 5 concludes the paper.

2. The Competence Paradigm & Competence Management

The competence approach to human resources management has a long history. The early Romans already practiced a sort of competence profiling in attempts to detail the attributes of a "good Roman soldier" (Draganidis & Mentzas, 2006, p.52). More recently, the concept of competence was used by early 20th century scientific management (Taylor, 1911) and later revised and redefined by McClelland, former Hay Group director, in the early seventies. According to McClelland (1973), competence concerns the relation between humans and work tasks: rather than knowledge and skills themselves, competence involves the knowledge and skills required to perform a specific job or task in an efficient way. From the sixties, and due to the ever-growing diffusion of the competence paradigm in business environments, the large number of proposed definitions made several authors claim they were facing a "competence pandemonium" (DeHaro, 2004). In a recent survey of Competence Management (CM), Draganidis & Mentzas (2006) state that a competence could be defined in terms of:

- *Category*. A group to which homogeneous and/or similar competences belong.
- *Name*. A descriptive name for the specific competence.
- *Definition*. Statement(s) that explains the basic concept of this competence.
- *Demonstrated behaviour*. Conduct indicators which demonstrate the acquisition of the specified competence.

According to the People Capability Maturity Model (Curtis, Hefley & Millar, 2001), competence management is a collection of employee management practices used to enhance the capability of the workforce to perform their tasks and to achieve the assigned competency growth objectives. CM is becoming more and more important in today's organizations. The employees' competence levels are important to achieve company goals and play a complementary role to, for instance, core business processes, customer relationships and financial issues (Norton & Kaplan, 1996). Based on a survey of CM systems, Draganidis & Mentzas (2006) point out that competences are capital in the following employee management applications: Workforce planning, Recruitment management, Learning management, Performance management, Career development and Succession planning.

Following the increasing importance of CM, an ever-increasing number of commercial Information Systems include CM modules. Based on a survey of commercial CM systems, Draganidis & Mentzas (2006) outline that open standards (XML, web services, RDF), semantic technologies (ontologies and the semantic web) and portals with self-service technologies represent the most extensively researched areas for the development of those applications. In particular, semantic technologies will be further discussed in the following sections.

The scientific literature reports several initiatives to implement Knowledge Management systems and, particularly, CM systems in software engineering (Rus, & Lindvall, 2002), (Lindgren, Stenmark, & Ljungberg, 2003), (Aurum, Parkin, & Cox, 2004), (Ward, & Aurum, 2004), (Dingsøyr, Djarraya & Røyrvik, 2005). Human Factors represent one of the most important areas of improvement in Software Engineering (SE). Boehm points out that "After product size, people factors have the strongest influence in determining the amount of effort required to develop a software product " (Boehm, 1981), and "Personnel attributes and Human Resource activities provide by far the largest source of opportunity for improving software development productivity" (Boehm, Horowitz, Madachy, Reifer, Clark, Steece, Brown, Chulani, & Abts, 2000). However, CM systems for Social Networks still represent an under-researched field which and thanks to the progressive globalization of software development companies constitutes an interesting development area, since they will offer the possibility of searching, locating and sharing human resources all around the world, wherever they are needed by a project, a company or an organization.

3. Semantics-based Social Networks

In this section, we will firstly describe social networks and how they can foster resource sharing. Secondly, we characterise the Description of a Project (DOAP) ontology and explain why it represents a well-known conceptual model for sharing and interlinking software projects. Social networks provide the capability of managing information about knowledge and skills owned by a set of individuals. However, when querying that information no semantic mechanisms can be used. The combination of both social networks and semantic web approaches outperforms the previous attempts of purely linking resources or fully-fledged social links, since they provide a meaning to such relationships.

3.1. Social Networks

Social network (Al Hasan, Chaoji, Salem, & Zaki, 2006) is a popular way to model the interactions among people in a group or community. A social network can be visualized as a graph where nodes correspond to individuals and edges represent the association among them. Associations are usually driven by mutual interests that are intrinsic to the examined group. Social networks are usually dynamic objects, since new edges and nodes add to the graph over time. Understanding the dynamics that drive the evolution of a social network is a complex problem, due to a large number of variable parameters. A comparatively easier problem is to understand what drives the association between two specific nodes. Several variations of the above problems provide interesting research topics. For instance, some interesting questions concern the change of association patterns over time, the factors shaping the associations, the impact of the association between two nodes over other nodes.

Natural examples of social networks include the community of scientists in a particular discipline, with edges joining pairs who have co-authored papers (Newman, 2001); the set of all employees in a large company, with edges joining pairs working on a common project; a collection of business leaders, with edges joining pairs who have served together on a corporate board of directors. The availability of large, detailed datasets encoding such networks has stimulated extensive study of their basic properties and the identification of current structural features.

The size of a social network is not a trivial calculation. However, most research on that direction builds up on graph theory. The most commonly addressed variables are average distance between reachable pairs, distance-based cohesion or the number of hops following the shortest path among two nodes of the network (Batagelj & Mrvar, 2003). The distance between the two most distant nodes can be envisaged as the biggest distance of the network. This calculation can be achieved using graph-oriented applications such as (Bategelj & Mrvar, 2003) or UCINET (Borgatti, Everett, & Freeman, 2002). These applications fundamentally input the social network as a "contact matrix" in a Comma Separated Value format, which describes the network structure and provides precise information about the aforementioned metrics.

Most real world social networks are described by the so-called Small World phenomenon, which is familiar to anyone who has said "It's a small world" upon discovering a mutual acquaintance with a stranger or meeting somebody in a plane. In the late sixties, social psychologist Stanley Milgram tested this phenomenon experimentally by asking a set of subjects in Omaha, Nebraska to deliver a message to a specific target in Boston, Massachusetts (Adamic & Adar, 2004). The participants could pass the message only to people they knew on a first name basis and yet the message was passed an average of six times only. This evidence, synthesised by the expression "six degrees of separation", is somehow surprising, especially when considering that most people tend to move in close social circles tied to a geographic location, profession or activity.

The Small World phenomenon was formalized by Watts & Strogatz (1998), who also proved that adding some specific links, the diameter of the social network dramatically decreased, whereas the clustering coefficient was not significantly affected. In addition, they showed that social networks can be framed as Small World networks. Subsequently, several studies showed that also the World Wide Web was also a Small World network. Small World networks present two interesting features. On the one hand, when increasing their size, the diameter increases very smoothly, namely at a logarithmic rate. On the other hand, most of the nodes are connected to each other through a reduced number of connectors (Gladwell, 2002).

In principle, all nodes in a social network should be the same. However, the reality is that some of them "are more the same than others". First of all, multiple nodes are connected via a small number of nodes that are in-between. This phenomenon, known as betweenness (Freeman, 1977), means that these nodes are acting as hubs of connections, playing a fundamental role in the Social Network (Gladwell, 2002). Following Gladwell, connectors are people in a community who know large numbers of people and who are in the habit of making introductions. A connector is essentially the social equivalent of a computer network hub. Connectors usually know people across an array of social, planes, cultural, professional, and economic circles, and make a habit of introducing people who work or live in different circles. Although connectors are rare – only one in several thousand people might be thought of as a true connector – they play a relevant role in the network since they channel the relationships between other nodes.

3.2. Semantic: The Description of a Project ontology

The term "Semantic Web" was coined by Berners-Lee, Hendler & Lassila (2001) to describe the evolution from a document-based web towards a new paradigm that includes data and information for computers to manipulate. Ontologies (Fensel, 2002) are the technological cornerstones of the Semantic Web, because they provide structured vocabularies that describe a formal specification of a shared conceptualization.

The fundamental aim of the Semantic Web is to answer the ever-growing need for data integration on the Web. The benefit of adding semantics consists of bridging nomenclature and terminological inconsistencies to include underlying meanings in a unified manner. Given that a universally shared data format is not likely to arise and diffuse, the Semantic Web provides an alternative solution to represent the comprehensive meaning of integrated information and promises to lead to efficient data managing by establishing a common understanding (Shadbolt, Hall, & Berners-Lee, 2006).

The *de facto* Semantic Web standard ontology language is OWL (Web Ontology Language)[1], a markup language for publishing and sharing data on the Internet using ontologies. A more lightweight ontology language is the Resource Description Framework (RDF)[2], a family of specifications for a metadata model that is often implemented as an application of XML. The RDF family of specifications is maintained by the World Wide Web Consortium (W3C). The RDF metadata model is based upon the idea of making statements about resources in the form of a subject-predicate-object expression, called a triple in RDF terminology. The subject is the resource to be described. The predicate is a quality of that resource, which often expresses a relationship between the subject and the object. The object is the content of the relationship or the value of the examined quality. The simplicity of data modelling and the ability to model disparate, abstract concepts led to the increasing diffusion of RDF also in knowledge management applications unrelated with Semantic Web.

The Description of A Project (DOAP)[3] ontology constitutes an effort to provide a RDF vocabulary to describe a software project (Figure 1) also based on a OWL syntax.

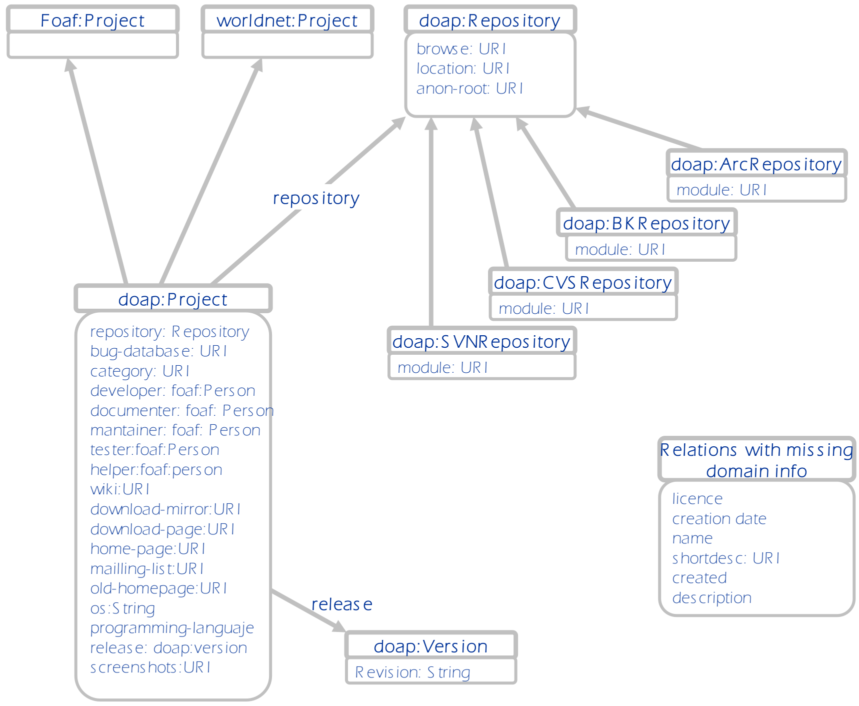

Figure 1. The DOAP Ontology

The DOAP ontology includes several relationships that can be used in the context of a software project and provide a meaningful description of the software modules of each project. Those features help tracking the evolution of the project, the sources of software (e.g., its the involved repositories), which developers participated in different projects and releases, etc.[4]

3.3. Competency & Semantics integration

Since the irruption of Semantic Web, there has been a significant amount of research and development initiatives focused on extending current web technology with machine-understandable metadata, aimed at providing layered services (Sicilia & Lytras, 2005). Semantic Web has been applied in multiple processes within the boundaries of what some scholars define as the Learning Organization (Örtenblad, 2001), a particular type of organization where learning behaviours improve and adapt, where individual and collective learning is encouraged and managers are coaches rather than directors. As argued by Sicilia & Lytras (2005), a number of potential applications of semantic technologies can be found in these organizations, ranging from Learning Environment, to Identifying Learning and Development Needs, Meeting learning and development needs and Applying learning in the workplace. Within these categories, authors present a fully-fledged set of Semantic Web Technologies whose impact is tremendous in a number of the aforementioned domains. Broadening the scope of the authors reference about competency and semantics integration, we can refer to ontologies of competence (Sicilia, 2005) and descriptions of knowledge management learning activities (Sicilia, Lytras, Rodríguez & García-Barriocanal, 2006).

One of the most relevant and ambitious projects, TenCompetence, is funded by the European Commission through the IST Programme. TenCompetence is developing an infrastructure to support individuals, groups and organisations in lifelong learning by integrating models and tools to create, store and exchange knowledge resources, learning activities, competence development programmes and network data for lifelong learning. Some of the researchers of TenCompetence have participated in the integration of a explicit competence ontology (Vasconcelos, Kimble & Rocha (2003), (Colucci, Di Noia, Di Sciascio, Donini, Mongiello, & Mottola, 2003), (Posea & Harzallah, 2004). More recently and in the same field of competence ontologies, Draganidis, Chamopoulou & Mentzas (2006) proposed an Ontology-based Tool for Competency Management and Learning Paths in the eLearning field.

In our approach, we provide a proof-of-concept architecture and implementation of a semantics-based social network that identifies competences through the well-founded semantics of the DOAP ontology. These semantics cover the competences by providing a technological infrastructure that can benefit from semantics forthcomings such as inference, reasoning, relationship identification and non a-priori envisaged relationships.

In the next section, we present the ProLink architecture and how it has an impact of the proof-of-concept approach described in this work.

4. ProLink: A Semantics-based Social Network for Software Projects

In this section, we present a novel and promising architecture that combines both semantics and social networks in the field of competence management. We propose a tailor-made value-adding technological solution which addresses the aforementioned challenges and solves the integration problem regarding to searching, finding, interacting and integrating of heterogeneous sources by means of semantic technologies.

4.1 Prolink Architecture

The ProLink architecture is composed of the components depicted in Figure 2.

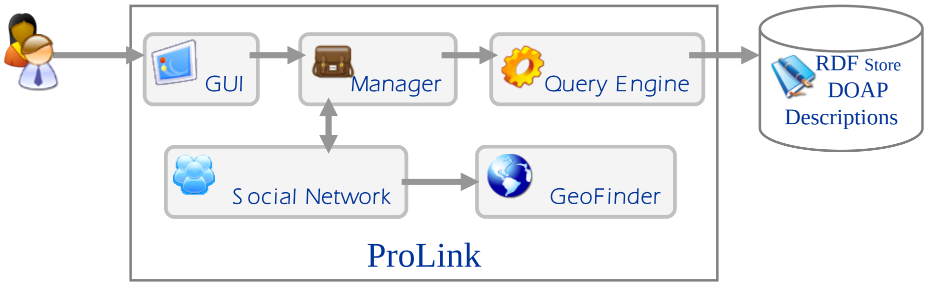

Figure 2. The ProLink Architecture

These components will be detailed in what follows.

- Social Network: The Social Network component is a set of integrated tools for semantically mapping the members of the ProLink social network that share a number of similarities in their projects through their DOAP descriptions.
- GeoFinder: It is a software agent which invokes the *PlaceFinder web service* [5] provided by ESRI to find the exact geographic location (longitude and latitude) of the places mentioned in the profiles of network members.
- RDF Repository: The RDF repository is a semantic data storage system that allows semantic querying and offers a higher abstraction layer to enable fast storage and retrieval of large amounts of RDF while keeping a small footprint and a lightweight architecture approach. An example of these systems could be the OpenRDF Sesame RDF Storage system[6] or the Yet Another RDF Storage System (YARS)[7], which deal with data and legacy integration. The advantages of using RDF as a "lightweight" ontology language partially rely on Faceted Search and Browsing techniques. These techniques will be analyzed at the end of the section.
- Query Engine: The Query Engine component uses a query language to make queries into the RDF storage system. The semantics of the query are defined not by a precise rendering of a formal syntax, but by an interpretation of the most suitable results of the query. Since the aforementioned systems store RDF triples (see section 4.2 for more details about Semantic Web languages), several query languages may be used. The SPARQL Query Language for RDF[8] is the most suitable query language. Since YARS enables SPARQL querying, due to pragmatic reasons this is the query language of our choice.
- Manager: Manager component is the main component of the architecture. It controls the different interactions among the components. Manager communicates with the query engine and the RDF repository to ensure that the information extracted from the user matches with the information stored in the DOAP, i.e. there is a match between the project the user wants to find and know more about and what is stored.
- GUI: This is the component that interacts with the user. It collects the request from the user and presents the obtained results. In our particular architecture, the GUI will collect requests pertaining to search criteria, such as, for example, "projects regarding a development in C++ of ERP systems". The GUI transfers the user request to the Manager component and displays the results the latter provides.

As outlined before, one of the main advantages of using RDF in the ProLink architecture is the possibility to use Faceted Search and Browsing. Faceted browsing is an information browsing technique consisting in navigating a dataset by partitioning the information space into orthogonal conceptual dimensions (Yee, Swearingen, Li & Hearst, 2003). These dimensions are called facets and represent important traits of the information elements. Each facet has multiple restriction values and the user selects a restriction value to constrain relevant items in the information space; in other words, each facet corresponds to the possible values of a property common to a set of digital objects. Faceted browsing is a visual query paradigm: the user constructs a selection query by browsing and adding constraints. Each step in the interface constitutes a step in the query build up and the user sees intermediate results and possible future steps while constructing the query. In ProLink, an exploratory interface allows users to find information navigating through the DOAP ontology data schema.

4.2 Example: Use of ProLink

This section illustrates the use of ProLink with a real-world case study scenario which shows the breakthroughs of our system. Let us consider a small company in Seattle, which is developing a typical three layer e-commerce system intended to perform the functionalities of an auction system. The software development team of the company has to make several choices regarding software requirements, architecture design, implementation language and so on. The access to a knowledge repository or a software portfolio to gather information about previous experiences on similar projects could usefully support those decision processes.

Fundamentally, the knowledge needs of the software development team concerns three main questions: did anyone already design and implement a similar system (for example, in Java)? Is it possible to access the related documentation? (Geographical, cultural and language factors have to be considered) Which other projects is their one related with? (there could be similar projects implemented in Python or Perl).

ProLink fully addresses these three knowledge needs. First of all, the software development team would use the GUI of ProLink to visually compose the most suitable information space query, for example, an English speaking, US-based, Java project. Given that ProLink uses faceted search and browsing, each requirement represents a facet or a trait which will divide the information space and will allow browsing all projects pertaining to such a space.

Thereafter, ProLink retrieves a number of related projects using the different components of the architecture. Let’s assume that the semantic descriptions and the Social Network component produce two results about two open-source development teams with similar projects. The first one is in Plano, Texas, and they have a lot of documentation in their portfolio, advising to use the Ruby on Rails (RoR) language for the implementation. The second team is Pittsburgh, PA, and also provides links to similar projects in the ProLink network using C++ programming language. The GeoFinder component allows showing the geographic location of both projects (Figure 3).

After having queried ProLink, the software development team would from then on be part of the ProLink network and share their experience. They would choose the most similar project and be granted access to its repositories, and all information pertaining the software project.

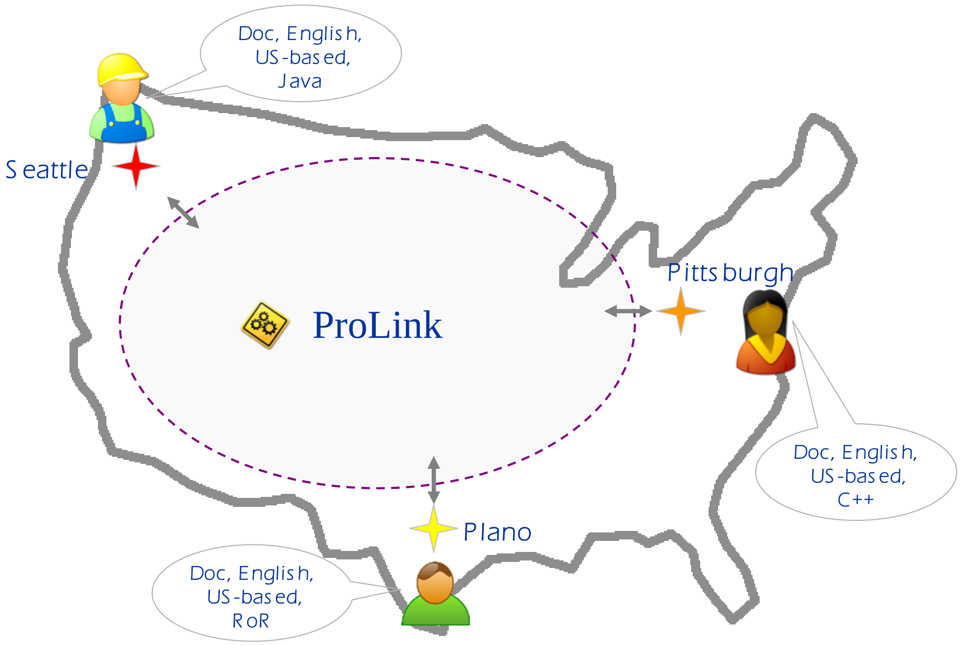

Figure 3. ProLink case study scenario

Actually, the Social Network component of the ProLink architecture complies with a typical social network, as described in section 3.1, but the ProLink social network at this stage focuses on main nodes able to address most of the requests. In the real world, those more developed and mature software organizations will have a remarkable portfolio and will be able to satisfy most requests.

The advantages of the connectors and betweenness approach in the ProLink social network are twofold. On the one hand, it takes into account dynamic trust at different nodes of the network, since connectors are highly reputed and trustworthy sources which can handle most requests. On the other hand, it also addresses the problems posed by the heterogeneity of network participants, since connectors are similar and are chosen on the basis of a reliable and efficiency-driven long term relationship.

Eventually, we would like to stress the advantages of the ProLink approach face to similar approaches. Clearly, the use of semantics and, more particularly, of a well-known and widely used ontology such as DOAP together with a clear conceptual competence framework has been a diamond in the rough throughout this work. In ProLink, the degree of expression of the semantic query languages has reached a mature state and it proves how a little semantics on a software project description can go a long way. Moreover, the use of advanced social networks concepts such as connectors and betweenness allows to take into account issues such as the reliability and trust of the network. Last but not least, the combination of social networks, semantics and software engineering represents a crucial issue for exploring new promising research lines and directions.

5. Conclusions

Integration of competences through semantics is a growing and recognized challenge that can revolutionize the IT working environment as we know it. With the rise of the Semantic Web, the ontology-based approach to social networks has gained momentum. In such a context, sharing and taking advantage of a number of information sources, tracing skills, experience and expertise in different fields can bridge the gap of knowledge integration. However, the ontology-based approach best works in specialized domains and environments where concepts and vocabularies can be well controlled.

In this work, we have presented a novel approach to achieve integration using the DOAP ontology to find the aforementioned value-adding features to our work through a semantics-based social network, ProLink, providing an architecture and a proof-of-concept implementation.

Our future work will focus on finding more user cases and real-world scenarios to determine the feasibility of our approach and validate its efficiency.

References

Adamic, L.A. & Adar, E. (2005) How to search a social network. Social Networks, 27(3), pp.187-203.

Al Hasan, M., Chaoji, V., Salem, S. & Zaki, M. (2006) Link Prediction using Supervised Learning. Proceedings of the SIAM Conference on Data Mining. Bethesda, Maryland.

Aurum, A., Parkin, P. & Cox, K. (2004). Knowledge Management in Software Engineering Education. Proceedings of IEEE International Conference on Advanced Learning Technologies, pp. 370 – 374.

Batagelj, V. & Mrvar, A. (2003). Pajek. Program for Large Network Analysis. University of Ljubljana, Slovenia, Ljubljana. http://vlado.fmf.uni-lj.si/pub/networks/pajek/doc/pajekman.htm

Berners-Lee, T., Hendler, J. & Lassila, O. (2001). The Semantic Web, Scientific American, May 2001.

Boehm, B. (1981). Software Engineering Economics. Prentice-Hall, Englewood Cliffs.

Boehm, B., Horowitz, E., Madachy, R., Reifer, D., Clark, B.K., Steece, B., Brown, A.W., Chulani, S. & Abts, C. (2000). Software Cost Estimation with COCOMO II, Prentice Hall, Upper Saddle River.

Borgatti, S. P., Everett, M. G. & Freeman, L. C. (2002), Ucinet 6 for Windows: Software for social network analysis. Harvard, Analytic Technologies.

Colucci, S., Di Noia, T., Di Sciascio, E., Donini, F. M., Mongiello, M. & Mottola, M. (2003). A formal approach to ontology-based semantic match of skills descriptions. J.UCS: The journal of universal computer science, 9(12), pp.1437 - 1454.

Curtis, B. Hefley, W.E. & Miller, S.A. (2001). People Capability Maturity Model. Version 2.0, CMU/SEI-2001-MM-01.

DeHaro, J.M. (2004). ¿Sabe alguien qué es una competencia? De McClelland a la ISO 9000. (Does anybody know what is competeny? From McClelland to ISO 9900) Revista de la Asociación Española de Dirección de Personal (Journal of Personal Management Spanish Association), 30, pp. 8-17.

Dingsøyr, T., Djarraya, H.K. & Røyrvik, E. (2005). Practical knowledge management tool use in a software consulting company. Communications of the ACM, 48(2). pp. 96-100.

Draganidis, F. & Mentzas, G. (2006). Competency based management: a review of systems and approaches. Journal: Information Management & Computer Security, 14(1), pp.51 – 64.

Draganidis, F. Chamopoulou, P. & Mentzas, G. (2006). An Ontology Based Tool for Competency Management and Learning Paths. International Conference on Knowledge Management.

Fensel, D. (2002) Ontologies: A Silver Bullet for Knowledge Management and Electronic Commerce. Springer-Verlag, Berlin.

Freeman, L. (1977). A set of measures of centrality based upon betweenness. Sociometry, 40(1), pp. 35–41.

Freeman, L. (1979). Centrality in social networks. Conceptual clarifications. Social Networks 1, pp. 215-239.

Gladwell, M. (2002). The Tipping Point. Back Bay Books, Boston.

Kretschmer, H. (1997). Patterns of behaviour in coauthorship networks of invisible colleges. Scientometrics, 40(3), pp.579–591.

Libel, D. & Kleinberg, J. (2003) The Link PredictionProblem for Social Networks. Proceedings of the 12^th ACM International Conference on Information and Knowledge Management (CIKM03).

Lindgren, R., Stenmark, D. & Ljungberg, J. (2003). Rethinking Competence Systems for Knowledge-Based Organizations. European Journal of Information Systems, 12(1), pp. 18-29.

McClelland, D. C. (1973). Testing for competence rather than for 'intelligence'. American Psychologist, 28, pp. 1–14.

McConnell, S. (2003). Professional Software Development. Addison-Wesley, Boston.

Newman, M. (2001).Who is the best connected scientist? A study of scientific coauthorship networks. Physics Review, 64(4).

Norton, D.P. & Kaplan, R.S. (1996). The balanced scorecard: translating strategy into action. Harvard Business School Press, Boston.

Örtenblad, A. (2001). On differences between organizational learning and learning organization, The Learning Organization, 8(3), pp. 125-33.

Posea, V. & Harzallah, M. (2004). Building a competence ontology. Proceedings of the workshop Enterprise modelling and Ontology of the International Conference on Practical Aspects of Knowledge Management (PAKM 2004) Wien, Austria.

Pressman R. S. (2005). Software Engineering: A Practitioners Approach (6th Edition), McGraw Hill, New York.

Rus, I. & Lindvall, M. (2002). Knowledge management in software engineering. IEEE Software, 19(3), pp. 26-38

Shadbolt, N., Hall, W. & Berners-Lee, T. (2006) The Semantic Web Revisited. IEEE Intelligent Systems 21(3), pp. 96-101.

Sicilia, M.A. & Miltiadis, L. D. (2005). The semantic learning organization. The Learning Organization, 12(5), pp. 402-410.

Sicilia, M.A. (2005). "Ontology-based competency management: infrastructures for the knowledge-intensive learning organization", in Lytras, M. and Naeve, A. (Eds), Intelligent Learning Infrastructures in Knowledge Intensive Organizations: A Semantic Web Perspective, IDEA, Hershey, PA.

Sicilia, M.A., Lytras, M., Rodríguez, E. & García-Barriocanal, E. (2006). Integrating descriptions of knowledge management learning activities into large ontological structures: a case study, Data and Knowledge Engineering, 57(2), pp.111-121.

Taylor, F. (1911) The Principles of Scientific Management. Harper & Row, New York.

Vasconcelos, J. B., Kimble, C. & Rocha A. (2003). Ontologies and the Dynamics of Organisational Environments. An example of a Group Memory System for the Management of Group Competencies. The 3rd International Conference on Knowledge Management, Graz, Austria.

Ward, J. & Aurum, A. (2004). Knowledge Management in Software Engineering: Describing the Process. Proceedings of the 2004 Australian Software Engineering Conference (ASWEC'04), pp.137-146.

Watts, D.J. & Strogatz, S.H. (1998). Collective dynamics of 'small-world' networks. Nature 393, pp. 440-442.

Yee, K. P, Swearingen, K., Li, K. & Hearst, M. (2003). Faceted metadata for image search and browsing. In Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI '03), Fort Lauderdale, FL, USA, April 5 – 10. New York: ACM Press pp. 401 – 408.

---

[1] OWL, http://www.w3.org/TR/owl-features/

[2] RDF, http://www.w3.org/RDF/
[3] DOAP, http://usefulinc.com/doap/

[4] As a side remark, *DOAP:project* is not a subclass of the *wordnet:project* since Wordnet is just a lexical thesaurus; the meaning of the relationship is just implying that DOAP is a project which defines a "semantic lightweight" vocabulary, same as the Friend of a Friend (FOAF) project. For example, we know that `doap:Repository` has a number of subclasses, such as `doap:ArchRepository`, `doap:BKRepository`, `doap:CVSRepository`, `doap:SVNRepository`, and the there are some relations that have a Repository as their domain (`doap:browse`, `doap:location`) and that the relation `doap:repository` has a `doap:Repository` as its range. This could help us to identify software repositories of this project, experiences gained, log of changes, etc.

[5] PlaceFinder, http://www.placefinder.com

[6] Open RDF Sesame, http://www.openrdf.org/

[7] YARS, http://sw.deri.ie/yars

[8] SPARQL, http://www.w3c.org/sparql