# Trust and Reputation in Social Internetworking Systems

Lora Aroyo[1], Pasquale De Meo[2], Domenico Ursino[2]

[1] Department of Computer Science, VU University Amsterdam, De Boelelaan 1081, 1081 HV Amsterdam, The Netherlands
[2] DIMET, Università "Mediterranea" di Reggio Calabria, Via Graziella, Località Feo di Vito, 89060 Reggio Calabria, Italy
l.m.aroyo@cs.vu.nl,demeo@unirc.it,ursino@unirc.it

**Abstract** A Social Internetworking System (*SIS*) is the scenario arising when Web users decide to affiliate to multiple social networks. Recent studies show an increasing user tendency of creating multiple identities on different social systems and exposing, in each of them, different traits of their personalities and tastes. This information provides a better picture of user needs and enhances the quality of services they can use. In the next years a large growth of *SIS* phenomenon is foreseeable. In order to boost the level of user participation in a *SIS*, suitable mechanisms capable of discerning reliable users must be designed. We propose a model to represent a *SIS*, a software architecture to gather real data and analyze the structural properties of a *SIS*. In concrete use cases with different contexts and different levels of protection of data, we introduce an ontology-based model to compute trust and reputation in a *SIS*. This research is collaborative effort between the Vrije Universiteit Amsterdam and the University of Reggio Calabria in the context of a Marie Curie Fellowship.

## 1 Introduction

Social media applications, such as blogs, multimedia and link sharing sites, question and answering systems, wikis and online forums, are growing at an unprecedented rate and are estimated to generate a significant amount of the contents currently available on the Web [14]. Social media applications are a significant part of a more meaningful kind of applications, named *Web 2.0 applications*, which aim to provide a platform for information sharing and collaboration among users on the Web.

In social media applications, users form communities, typically modelled as *social networks*. Users are driven to get in touch and become friends of other users, create and publish their own contents (like videos or photos), share these contents with others, rate and comment contents posted by others. Examples of popular Web-based social networks are Facebook, MySpace and LinkedIn.

The value of social networks expresses in multiple ways. For instance, users may take an advantage of their interactions with other users to find information

relevant to them or they can explore connections existing in a social network to get in touch with user with whom they may profitably interact: many Web users, as an example, indicate that they were able to get a job through their contacts in LinkedIn[1].

A further advantage is that social networks allow to disseminate new knowledge in a widespread fashion, to diffuse innovations, to spread opinions (e.g., social or political messages) among members, to advertise new products [16].

The power of social networks has been fully recognized by institutional actors like museums, TV broadcaster, academic and government institutions. For instance, the Rijksmuseum Amsterdam is exploring the added-value of providing artworks online, allowing users to express their opinions on them or contribute to describing artwork's. Furthermore, major European broadcasters, such as BBC and RAI are experimenting with Web 2.0 technologies to improve interactivity and participation of TV consumers.

Users often decide to affiliate to multiple social networks: for instance, in a recent survey, Ofcom found that 39% of UK adults with at least one social networking profile has indeed two or more profiles[2]. We call *Social Internetworking System* (hereafter $SIS$) the scenario arising when many users decide to affiliate with multiple social networks. Companies are discovering the potential of social internetworking and are promoting systems capable of supporting social internetworking tasks. For instance, Google has recently proposed Open Social [4], a set of APIs to access social sites, like LinkedIn or MySpace. Some systems (like FriendFeed [2]) allow the users to share their activities with other users in multiple social networks; Gathera [3], also provide the users with a single interface to handle their accounts on multiple social networks.

The main goal of these systems is to offer a technological platform to ensure data portability among different social networks. The major bottleneck for the success of a $SIS$ is the absence of mechanism that helps users in finding other "reliable" users with whom they can profitably interact and discloses the presence of malicious users or spammers.

In the past, significant research efforts have been done to define and handle trust and reputation, as a large body of literature highlights [1,11,12,15,19,22].

However, in our opinion, several reasons explain a further investigation. A first research question is to provide a model capable of representing a $SIS$, its components and their relationships. In addition, it is necessary to gather real data about a $SIS$ in order to understand its structural properties and clarify to what extent a $SIS$ differs from traditional social networks.

A second issue depends on the fact that the concepts of trust and reputation may assume different meanings according to the scenario in which a user operates in. For instance, in communities like Question & Answering systems (in which users are allowed to pose questions, to answer questions raised by other users

[1] `http://www.mainstreet.com/article/career/employment/social-media-job-seeker-s-best-new-tool`

[2] `http://www.ofcom.org.uk/advice/media_literacy/medlitpub/medlitpubrss/socialnetworking/annex3.pdf`

and, finally to rate received answers), the reputation of a user coincides with his level of expertise on a particular topic; in a Web community like YouTube, the reputation a user coincides with the quality of contents he generated. This requires to define a procedure to compute, in an abstract and general fashion, the reputation of a user and to specialize it in concrete domains.

As a final research challenge, it would be necessary to define a model to represent trust and reputation in different contexts. In addition, it is useful to observe that, in different contexts, different policies for accessing, publishing and re-distributing data may exist. For instance, in the case of a TV broadcaster which delivers online part of its archive of resources, users are allowed for instance to use some resources (e.g., for educational purposes) but are forbidden to re-use those protected by copyright. To address these issues, it would be beneficial to design an ontology to model the key concepts of trust and reputation in different environments characterized by different levels of protection of data.

In this paper we propose a methodology to handle trust and reputation in a $SIS$. The paper summarizes the research activities we are planning to carry out in the context of a *Marie Curie Intra-European Fellowships for Career Development (IEF)*, a funding opportunity provided by EU Commission. The paper is structured as follows: in Section 2 we review existing trust and reputation models and illustrate the challenges arising in a $SIS$. In Section 3 we illustrate a model to represent the features of a $SIS$ along with a software architecture we are currently implementing to gather real data from a $SIS$ and analyze its structural properties. In Section 4 we provide a general model to compute trust and reputation in a $SIS$ and illustrate the steps we are planning to specialize it in real contexts; in particular, we plan to adapt our notion of trust on data gathered within two research projects, namely *NoTube* [18] and *Agora* [5]. In Section 5 we discuss a possible ontology-based model to represent trust and reputation in environments characterized by different levels of protection of data. Finally, in Section 6 we draw our conclusions.

## 2 Background and challenging issues

In virtual communities the term *trust* is generally exploited to indicate the reliance that a community member associates with another one. Trust values are "local parameters" in the sense that specifying the trust of a user $A$ toward a user $B$ is equivalent to indicate how much $A$ perceives $B$ as reliable.

The opinion of the whole community of users toward a member of the community itself is known as *reputation*. In the past, the issue of computing and handling trust and reputation in virtual communities has been deeply investigated and several models and approaches to facing it have been proposed.

Here we discuss some of these approaches and outline the challenges we encounter in the context of a $SIS$. Existing approaches can be classified into two categories:

**Graph-Based Approaches**. A first category of approaches model a user community as a graph $G$ in which nodes represent users [1,11,12,22]. An edge

linking two nodes $v$ and $u$ indicates that the user $v$ explicitly trusts the user $u$. The graph $G$ is usually sparse because a user typically evaluates a handful of other users; as a consequence, various techniques have been proposed to *infer* implicit trust relationships. In detail, the approach of [1] applies a maximum network flow algorithm on $G$ to compute trust between any pair of users. In [11] the authors apply a modified version of the Breadth First Search algorithm on $G$ to infer multiple values of reputation for each user; these values are then aggregated by applying a voting algorithm to produce a final (and unique) value of reputation for each user. The approach of [12] considers paths up to a fixed length $k$ in $G$ and propagates the explicit trust values on them to obtain the implicit ones. In [22] trust values are computed by applying a spreading activation algorithm.

Graph-based approaches leverage on explicit trust relationships declared between pairs of users. As a consequence, they neglect to consider a broad range of activities that, in a $SIS$ (e.g., the activity of rating resources) are a precious and reliable indicator of trust.

**Link-Based approaches**. A second category of approaches use ranking algorithms such as PageRank [8] or HITS [17], which have been successfully applied in the context of Web Search, to find trust values. For instance, [15] proposes an approach based on PageRank to measure peer reputation in a peer-to-peer network. The approach of [19] defines a probabilistic model of trust which strongly resembles that described in [15]; however, differently from this last, the approach of [19] computes and handles trust values and not reputation values. In [9] the authors present an algorithm which computes global reputation values in a peer-to-peer network; the proposed algorithm uses a personalized version of PageRank along with information about the past experiences of peers.

Experimental tests indicated that link-based methods can obtain precise results and are often *attack-resistant*, i.e., they can resist to attempts conceived to manipulate reputation scores.

We observe that in some approaches trust is conceived as a *measure of performance*. For instance, in [15], the trust of a peer depends on the success of downloading a file from it and, then, trust depends on parameters like the number of corrupted files stored in the peer or the number of connections with the peer that have been lost. By contrast, in our case, trust should quantitatively encode the confidence of a user in the opinions formulated by other ones.

We can observe that both graph-based and link-based approaches try to model trust and reputation in a "force-mass-acceleration" style. In other words, these approaches try to capture all factors influencing trust and reputation and combine them in a set of equations. The resulting model is too complicated to be handled and it may not provide significant results. In our opinion, the assessment of trust and reputation critically depends on the concrete domain in which we are operating in and we believe that an universal model of trust is not possible.

To better clarify this concept, we report some results emerging from the *PrestoPrime* project [7], an EU financed project devoted to study and develop

practical solutions for the long-term preservation of digital media objects, programmes and collections.

In the context of PrestoPrime, two pilot demonstrators were developed. In the first one, in a game-like environment, users were asked to label videos by applying simple keywords (*tags*). Experiments with users showed that a satisfactory measure of trust between a pair of users who do not know each other can be obtained by considering the tags they apply to label a video and computing the degree of match of the set of tags they inserted.

In the second demonstrator, users were provided with a small annotation environment allowing them to label museum objects with four main entry fields (i.e., "who", "where", "what", and "when"). This allowed us to create links between museum objects on the basis of the key dimensions "who", "where", "what", and "when"; as an example, objects coming from different museum collections can be tied if they refer to the same artistic and historical context and this produces a more complete description of cultural movements.

The notion of trust developed in the context of the first demonstrator is not applicable for the second one, and other factors influencing trust and reputation need to be studied.

A further challenge we are in charge of studying depends on the fact that, in some cases, real organizations often decide to make available on the Web their own resources and often allow end users to enrich their descriptions through metadata like tags. For instance, think of the case of public TV broadcasters like BBC which offers online a large number of contents referring to its TV programmes. Each organization may use different policies for accessing, distributing and labelling the contents they produce and disseminate. For instance, a digital content may published online only in some specific cases (e.g., if the material must be used in education) while its usage is forbidden for commercial purposes. This proves that, in the process of defining trust and reputation, it is necessary to consider not only the application context but also te level of data protection about available resources.

## 3 Defining a basic model of social internetworking

The first goal of our research is to find a suitable model to represent a Social Internetworking System ($SIS$) and interactions between humans that can take place in it.

To this purpose, our model must fit two requirements:

- *Requirement 1.* The model should be rich enough to represent a wide range of *heterogeneous entities* (i.e,. users, resources, posts, comments, ratings, and so on) and their *interactions* (e.g., users may declare to be friends or they may rate resources).
- *Requirement 2.* The model should be easy to manipulate and intuitive.

Clearly, Requirements 1 and 2 are conflicting each other and a suitable trade-off is compulsory. Traditional approaches to modelling social networks are usually

based on *graphs*. Nodes in graphs represent social network actors (e.g., users) while edges identify relationships between them.

We believe that graph-based models are not satisfactory in the context of a $SIS$ for several reasons.

A first weakness relies on the role the nodes would have if we would decide to represent a $SIS$ through a graph. Generally, a social network consists of *homogeneous* nodes, i.e., all nodes represent objects sharing the same nature. As claimed by Requirement 1, in a $SIS$ *heterogeneous* entities co-exist and these heterogeneities must be properly modelled.

A further limitation is that graph-based models are able to represent *one-dimensional networks*, i.e., edges of a graph specify that only *one* particular kind of relationship may exist between nodes. On the contrary, we expect that a $SIS$ should be represented through a *multi-dimensional* network because various type of interactions may involve entities of the same type or of different nature: for instance, an edge should link a user $u$ and a resource $r$ to indicate that $u$ has posted $r$ or an edge should tie two users to indicate that they declared to be friends.

Finally, edges in graphs highlight *binary relationships* between nodes they link. In a $SIS$, it could be useful to consider $n$-ary relationships (e.g., an edge may glue together a user $u$, a resource $r$ and a tag $t$ under the hypothesis that $u$ applied $t$ to label $r$).

We are currently studying a more sophisticated model in which a $SIS$ is represented through an *hypergraph* such that: *(i)* nodes are labelled and the label of a node reflects the nature of the object represented by the node itself; *(ii)* multiple hyperedges may run between two nodes to indicate that multiple interactions may take place between two arbitrary entities; *(iii)* hyperedges denote relationships involving two or more entities.

In addition to defining a model to represent a $SIS$, we are also interested in gathering data from real social networks in order to understand the properties showed by a real $SIS$. For instance, it would be interesting to check whether properties typical of real social networks (e.g., the small world phenomenon) still emerge in a $SIS$.

Such a task is quite complex because, in different networks, a user may have different identities so it would be extremely hard to join information scattered across multiple networks.

To address this issue, we used the *Google Social Graph API* [6]. Social Graph API allows human users or software applications to access public connections between people on the Web. In particular, Social Graph API can be queried through an HTTP request and is able to return two kind of results:

- *A list of public URLs that are associated with a person.* For instance, given a user $u$, Social Graph API reveals the URLs of the blog of $u$ and his Twitter page.
- *A list of public declared connections between people.* For instance, it returns the list of persons who, in at least one social network, have a link to the blog of $u$ or any other page referable to $u$.

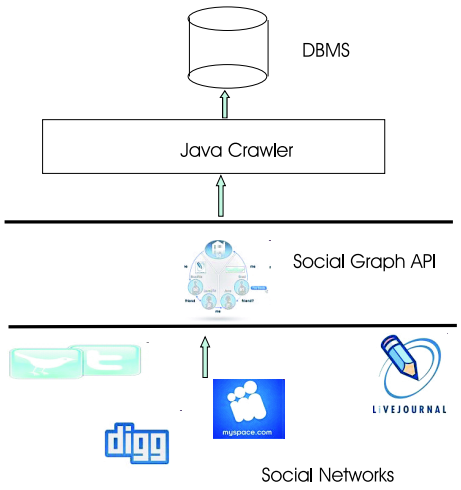


**Figure 1.** The architecture of our crawler

At the moment of writing, we have implemented a simple crawler to explore and gather data from a *SIS*. The architecture of our system is shown in Figure 1. In particular, a Java crawler invokes Social Graph API by sending a *seed URL* associated with a user $u$; the API sends back the list of people who are somewhat related with $u$. The Java crawler gets these data and launches a Breadth-First-Search like procedure to find new URLs and connections. Retrieved results are permanently stored in a relational DBMS implemented in MySQL.

A key issue in gathering user data is *privacy*. In fact, Social Graph API handles and returns only URLs of *public pages* and *publicly declared connections* between them. The API is not able to reveal non-public information (e.g., private profile pages or Web pages accessible to a restricted circle o persons). Despite these limitations, users may unintentionally leave digital traces of their interactions and, by putting together these traces, confidential data may be disclosed.

As a consequence, informing users about privacy risks is not enough to avoid privacy violations, but complex algorithms are required.

Traditional approaches to protecting privacy (e.g., those based on $k$-anonymity principle) have been proved to not be effective in the context of social networks [13]. With regard to this, we point out that some successful approaches (relying on the idea of carrying out a sequence of random insertions and deletion of edges in the social network graph) have been proposed [13] and, despite privacy is not the main focus of our work, we plan to study whether these approaches can be extended to a *SIS*.

## 4 Defining reputation in a Social Internetworking System

As a further step, we are interested in studying a model of reputation for a *SIS*. The notion of trust/reputation in the context of a *SIS* (and, in general, for Web

2.0 applications) is hard to define because, as shown in Section 2, in different contexts, they may assume different meanings.

We propose a methodology to compute reputation in a $SIS$ which operate in *two* stages: in the former stage we review and analyze the factors capable of influencing the value of reputation in a $SIS$. In the latter stage we consider some concrete domains and specialize our methodology to them.

Intuitively, we assume that the reputation of a user depends on the following facts:

*The reputation of a user depends on the relationship he created in the $SIS$.* We suggest to use the hypergraph model introduced in Section 3 to represent users and their relationships in a $SIS$. Past user interactions are analyzed to determine the level of trust a user $u$ confers to a user $v$ and this information is used to weight edges in the hypergraph representing the $SIS$. Since a user generally interact with an handful of other users, the hypergraph we obtain is *sparse* and a suitable algorithm to propagate trust values is necessary. Currently, we are planning to use a link-based algorithm like PageRank. At the end of this step, we are able to generate a vector $\mathbf{r}'$ such that the $i$-th component of $\mathbf{r}'$ equals to the reputation of the $i$-th user.

*Users with high level of reputation are also those who produce high quality resources.* The quality of a resource could be computed by consider the average rating it got and, then, resources with a high average rating are also high quality resources. To avoid biases, we can pose a further requirement: the number of ratings received by a resource must be *statistically significant*, i.e., we can consider only resources which received at least $N_{min}$ ratings, being $N_{min}$ a suitable threshold. Such a requirement would avoid that resources evaluated by a small number of users are deemed better than resources rated by a large mass of human users.

The procedure described above resembles that applied in many social systems like YouTube or Digg to evaluate the quality of a resource. We believe that such a procedure is affected by several fallacies and it may incur in harsh inaccuracies. In fact, spam or malicious users may tend to provide generous evaluations to artificially inflate the evaluation of a resource. As a consequence, we need a more complicated framework capable of putting together the reputation of users, the quality of resources they post and the evaluations associated with resources. At the current stage of the project we are considering, as a possible solution, the following criterium:

> *A user has a high reputation if he authors high quality resources. A resource, in its turn, is of high quality if it gets a high average rating and it has been posted by users with high reputation.*

The intuition provided above relies on a *mutual reinforcement principle* that is similar, to some extent, the approach underlying HITS [17] algorithm. The principle outlined above easily turns into a set of linear equations. In fact, let $n$ be the number of users composing a $SIS$ and let $m$ be the number of resources they authored. Let $\mathbf{r}''$ be an $n$-th dimensional array such that the $i$-th entry

of $\mathbf{r}''$ equals to the reputation (to compute) of the $i$-th user and let $\mathbf{q}$ be an $m$-th dimensional array such that the $j$-th entry of $\mathbf{q}$ equals to the quality (to compute) of the $j$-th resource. Finally, let $\mathbf{e}$ be an $m$-th dimensional array such that the $j$-th entry of $\mathbf{e}$ equals to the average rating of the $j$-th resource and let $\mathbf{A}$ be an $n$-by-$m$ matrix such that $\mathbf{A}_{ij}$ equals 1 if the $i$-th user posted the $j$-th resource and 0 otherwise.

According to this notation, we can write the following equations:

$$\mathbf{r}'' \propto \mathbf{A}\mathbf{q} \tag{1}$$

$$\mathbf{q} \propto \mathbf{A}^T\mathbf{r}'' + \mathbf{e} \tag{2}$$

In both Equations 1 and 2, the symbol $\propto$ means "is proportional to". As for Equation 1, the $i$-th row of the the product $\mathbf{Aq}$ specifies the sum of the qualities of the resources authored by the $i$-th user. This immediately follows from the definition of product between a matrix and a vector. Interestingly enough, the $\mathbf{A}$ matrix can be interpreted as the adjacency matrix of a bipartite graph whose nodes represent users and resources and edges link a user to the resources he authored. In Equation 2, the symbol $\mathbf{A}^T$ is the *transpose* of $\mathbf{A}$. As in the previous case, $\mathbf{A}^T$ matrix can be viewed as the adjacency matrix of a bipartite graph whose nodes represent resources and users and edges link a resource to the user who authored it. Observe that the same model holds if we assume that a resource has been posted by one user or it has been posted by multiple users. The product $\mathbf{A}^T\mathbf{r}''$ is an $m$-th dimensional vector whose $j$-th entry specifies the reputation (or the sum of the reputations) of the user (users) who posted the $j$-th resource.

By plugging Equation 2 into Equation 1 we obtain:

$$\mathbf{r}'' \propto \mathbf{A}\left[\mathbf{A}^T\mathbf{r}'' + \mathbf{e}\right] \Rightarrow \mathbf{r}'' - \mathbf{A}\mathbf{A}^T\mathbf{r}'' \propto \mathbf{A}\mathbf{e} \Rightarrow$$

$$\Rightarrow \mathbf{r}''\left[\mathbf{I} - \mathbf{A}\mathbf{A}^T\right] \propto \mathbf{A}\mathbf{e} \Rightarrow \mathbf{r}'' \propto \left[\mathbf{I} - \mathbf{A}\mathbf{A}^T\right]^{-1}\mathbf{A}\mathbf{e}$$

Since $\mathbf{A}\mathbf{A}^T$ is *symmetric*, its eigenvalues are real [20]. In particular, $\left[\mathbf{I} - \mathbf{A}\mathbf{A}^T\right]^{-1}$ can be easily and effectively approximated by computing the *dominant eigenvector* of $\mathbf{A}\mathbf{A}^T$. Such a result is of great practical impact because there exist efficient numerical methods to compute dominant eigenvector of a symmetric matrix (think of Lanczos method [20]) and, then, our methodology is suitable also if the size of $\mathbf{A}$ gets very large; such a case if quite common in real cases because, in traditional social sites the number of users and resources they generate (which correspond to the number of rows and columns of $\mathbf{A}$) is huge.

Finally, we merge the arrays $\mathbf{r}'$ and $\mathbf{r}''$ into a single reputation value $\mathbf{r}$ as follows:

$$\mathbf{r} = \alpha\mathbf{r}' + (1-\alpha)\,\mathbf{r}'' \tag{3}$$

The coefficient $\alpha \in [0,1]$ is instrumental in weighting the contributions coming from link analysis and the analysis of resources generated by a user. We plan to tune $\alpha$ by applying a linear regression technique.

Once a theoretical model of reputation in a *SIS* has been defined, our intention is to specialize it in concrete domains. In particular, we are interested in monitoring and analyzing the behaviour of users in long-term experiments associated with different domains; the notion of reputation, from abstract concept turns into a concrete tool to aid user in better taking advantage of potentialities offered by the *SIS*. Experiments on real users allow us to get an iterative assessment of the strengths and weaknesses of our notion of reputation as well as indications for improvement.

To this purpose, we will use data gathered in the context of two research projects, namely: *(i) NoTube* (an EU financed project on interactive television) [18], and *(ii) Agora* [5] (a Dutch funded project on digital museums and audiovisual archives). In the context of interactive television, trust/reputation values represent the level of expertise of a user. This information could be exploited to select in a personalized fashion contents to propose to the user. As for digital museums we can study what parameters in the user behaviour are relevant for producing authoritative annotations and what are the motivations for users to participate in this labelling process. This information could be instrumental in better using human mass potential in annotating artworks.

## 5 Building an ontology-based model of trust and reputation in a Social Internetworking System

Once we have defined the concept of trust and reputation in concrete domains it is advantageous to create a model capable of representing them in different domains. To this purpose we plan to design an *ontology* capable of specifying how reputation and trust specialize in different application contexts.

To the better of our knowledge, there are few approaches to designing ontologies to model trust. For instance, in [10], *TrustOntology* is presented. This is an OWL ontology allowing each user to indicate the people he trusts. Trust information is automatically composed to infer new values of trust for newcomer users. In [21], the authors suggest a trust protocol in which the decision about the trustworthiness of a message depends on many factors like the creator (*who*) of the message (*what*), time (*when*), location (*where*), and intent (*why*). An ontology to capture factors influencing trust and a set of functions to evaluate trust is presented.

Our goal is different from that of [10] and [21] we want to model how reputation specializes in different contexts. In addition our ontology can be used to represent a scenario in which different organizations decide to make available on the Web their own resources. An organization (e.g., a cultural institution) may let its users to freely use, copy and re-distribute available resources. Another organization may apply different policies to protect data because some resources can be freely disseminated, other resources are not accessible because protected by copyright and, finally, some resources can be published online and re-used for some purposes (e.g., as educational material) but their access is forbidden in other cases.

The availability of such an ontology would offer us, on the long run, the possibility of designing complex software applications running across multiple social networks. For instance, we can think of a content-based recommender system operating as follows:

1. A user issues a query.
2. The query is forwarded to multiple social system and a list of resources matching the query is retrieved by each social system.
3. Retrieved resources are ranked on the basis of the reputation of the users who created, on the application context and on the rights for its distribution.
4. A global list is produced by merging the previous ones.

Such an application is, in our opinion, capable of introducing relevant novelties in the research field of Recommender Systems. In fact, the proposed application is able to sift through different social sites (while traditional Recommender Systems usually operate on a single resource repository) and is able to rank resources on the basis of multiple and criteria.

## 6 Conclusions

In this paper we introduce the concept of Social Internetworking System, i.e., the scenario arising when Web users decide to affiliate to multiple social networks. We propose a model to represent a *SIS* and describe the main components of a software architecture we are implementing to gather real data from a *SIS* and analyze its structural properties. In concrete use cases with different contexts and different levels of protection of data, we introduced an ontology-based model to compute trust and reputation in a *SIS*. This research is collaborative effort between the Vrije Universiteit Amsterdam and the University of Reggio Calabria in the context of a Marie Curie Fellowship.

In the future we plan to gather a large amount of data about a *SIS* and carry out an empirical study on them. The goal is to understand whether some properties of real social networks (like small world phenomenon, power law distribution of in-degree and out-degree distributions, and so on) if they are still confirmed in a *SIS* or if significant deviations emerge.

A further research line is to carry out a detailed review of existing literature on the meaning of trust and reputation in different social site. Finally, we plan to test the effectiveness of our ontology-based model with an experiment involving real users. In particular, the validation phase will be strictly tied to the activity of designing our ontology; in fact, we shall use feedbacks provided by users to revise the structure of our ontology.

## References

1. Advogato's trust metric. `http://www.advogato.org/trust-metric.html`, 2000.
2. FriendFeed. `http://friendfeed.com/`, 2009.

3. Gathera. http://www.gathera.com/, 2009.
4. Google Open Social. http://code.google.com/intl/it-IT/apis/opensocial/, 2009.
5. Agora: Creating the historic fabric for and providing web-enabled access to objects in dynamic historical sequences. http://agora.cs.vu.nl/, 2010.
6. Google Social Graph API. http://code.google.com/intl/it-IT/apis/socialgraph/, 2010.
7. Prestoprime: Keeping audiovisual contents alive. http://www.prestoprime.org/, 2010.
8. S. Brin and L. Page. The anatomy of a large-scale hypertextual web search engine. *Computer Networks*, 30(1-7):107–117, 1998.
9. P. Chirita, W. Nejdl, M. T. Schlosser, and O. Scurtu. Personalized Reputation Management in P2P Networks. In *Proc. of the ISWC Workshop on Trust, Security, and Reputation on the Semantic Web*, CEUR Workshop Proceedings, Hiroshima, Japan, 2004. CEUR-WS.org.
10. J. Golbeck. Trust ontology. http://www.schemaweb.info/schema/SchemaDetails.aspx?id=171, 2010.
11. J. Golbeck and J.A. Hendler. Inferring binary trust relationships in web-based social networks. *ACM Transactions on Internet Technology*, 6(4):497–529, 2006.
12. R. Guha, R. Kumar, P. Raghavan, and A. Tomkins. Propagation of trust and distrust. In *Proc. of the International Conference on World Wide Web (WWW '04)*, pages 403–412, New York, NY, USA, 2004. ACM.
13. M. Hay, G. Miklau, D. Jensen, D.F. Towsley, and P. Weis. Resisting structural re-identification in anonymized social networks. *Proceedings of the VLDB Endowment*, 1(1):102–114, 2008.
14. P. Heymann, G. Koutrika, and H. Garcia-Molina. Can Social Bookmarks Improve Web Search? In *Proc. of International Conference on Web Search and Data Mining (WSDM 2008)*, pages 195–206, Stanford, California, USA, 2008. ACM Press.
15. S.D. Kamvar, M.T. Schlosser, and H. Garcia-Molina. The Eigentrust algorithm for reputation management in P2P networks. In *Proc. of the International Conference on World Wide Web (WWW 2003)*, pages 640–651, Budapest, Hungary, 2003. ACM Press.
16. J. Kleinberg. The convergence of social and technological networks. *Communications of the ACM*, 51(11):66–72, 2008.
17. J. M. Kleinberg. Authoritative sources in a hyperlinked environment. *Journal of the ACM*, 46(5):604–632, 1999.
18. L. Nixon and L. Aroyo. NoTube – Making TV a Medium for Personalized Interaction. In *Proc. the European Interactive TV Conference (EuroITV 2009)*, pages 22–25, Leuven, Belgium, 2009. University of Leuven.
19. M. Richardson, R. Agrawal, and P. Domingos. Trust Management for the Semantic Web. In *Proc. of International Conference on Semantic Web (ISWC 2003)*, pages 351–368, Sanibel Island, FL, USA, 2003. Lecture Notes in Computer Science, Springer.
20. G.W. Stewart. *Matrix Algorithms: Basic Decompositions (Volume 1)*. Society for Industrial Mathematics, 1998.
21. S. Toivonen and G. Denker. The Impact of Context on the Trustworthiness of Communication: An Ontological Approach. In *Proc. of the ISWC Workshop on Trust, Security, and Reputation on the Semantic Web*, volume 127 of *CEUR Workshop Proceedings*. CEUR-WS.org, 2004.
22. C. Ziegler and G. Lausen. Propagation models for trust and distrust in social networks. *Information Systems Frontiers*, 7(4-5):337–358, 2005.