

Towards Semantic WikiCrimes

Vasco Furtado^{1,3}, Leonardo Ayres¹, Marcos de Oliveira², Carlos Gustavo¹, Juliana Oliveira¹

¹ Universidade de Fortaleza – UNIFOR

Av. Washington Soares, 1321 Fortaleza-CE Brazil

² University of Otago

PO Box 56, Dunedin, 9054, New Zealand

³ ETICE – Empresa de Tecnologia da Informação do Ceará

Av. Pontes Vieira, 220 Fortaleza-CE Brazil

Abstract

Collaboration is on the rise, primarily leveraged by the Web 2.0 in which the quantitative difference between information producer and consumer decreases significantly. In this paper, we describe WikiCrimes, a typical Web 2.0 application that offers a collaborative environment, based on the use and direct manipulation of maps in order to register and search criminal events. WikiCrimes is driven by three goals: i) to give more transparency and publicity to criminal information, ii) to provide means for citizen prevention, and iii) to reduce the phenomena of under-reporting (crimes that are not reported to authorities). We pay particular attention to how WikiCrimes can be leveraged by semantic web by means of the use of explanation ontologies to connect people, provide information about reliable and reputable sources, and provide rich filter mechanisms for citizen information and prevention.

Keywords: Web 2.0, Collaborative Systems, Law Enforcement, Reputation, PML2, Ontology

Introduction

Despite the huge success of Wikipedia (<http://www.wikipedia.org>), finding good and useful causes which are capable of involving thousands or even millions of people operating with little or no coordination is still a big challenge. In the context of public and government areas, particularly in law enforcement, one of the difficulties is finding the equilibrium between people's participation and information credibility. Anonymous mass collaboration is the easiest way to receive information; however, the credibility of the information received is depreciated, because the source of information is unknown.

It is in that particularly complex domain that we have initiated a project called WikiCrimes (<http://www.wikicrimes.org>), which is driven by three goals: i) to give more transparency and publicity to criminal information, ii) to provide means for citizen prevention, and iii) to reduce the phenomena of under-

reporting (crimes that are not reported to authorities). These goals have been in the political agenda of several countries around the world, particularly, those in which the populace suffers with high rates of violence.

WikiCrimes aims to offer a common interaction space among the public in general, so that they are able to report criminal facts as well keep track of the locations where such crimes occur. The goal is to obtain collaborative individual participation for generating useful information for everyone. Here, we are appealing to the sharing feeling that is so usual in victims of violence. When someone is a victim of any type of crime, it is usual to tell someone about the fact. Typically, those who are close to the victim are the first to be informed. What we intend with WikiCrimes is to provide an environment to be that “global notebook” of stories told by people about crimes, in order to help them alert other people on a scale larger than their closest social contacts.

In this article we briefly describe the goals, motivations and main features of WikiCrimes in order to introduce our actual research, which intends to introduce semantic representation in WikiCrimes elements. We have modeled WikiCrimes as an Open Multiagent System in which ontologies play a fundamental role in the agent interactions. The description of our approach for putting semantics in WikiCrimes is driven by the description of the ontology representation. We describe how the PML2 explanation ontology [8], involving the concept of trust and reputation, can be linked with the crime and report the crime ontology, providing semantics to one of the key success factors of the project: the reliability of information. We then exemplify how data integration provided by the use of semantic web representation languages leads to a better exploration of WikiCrimes data from different contexts, fostering people's connectivity as well.

WikiCrimes: Motivation and Goals

The veracity and accuracy of information about where crimes occur, as well as the information on the characterization of such crimes, has always been on the agenda of discussions on Public Safety in Brazil and in

various other countries. Traditionally, this information is monopolized by law enforcement agencies and is therefore characterized as a highly centralized mechanism. This monopoly ultimately creates tension in the relationship between such agencies and society at large, because it is commonly opposed to the precept of disclosure and transparency of information required by a democratic regime. Allied to this context are the crises that have characterized the daily routines of law enforcement agencies, as well as their limitations to provide a quality public service, which tend to diminish citizens' trust in those agencies. These factors encompass some of the reasons for the growing problem of under-reporting—the low rate of reporting crimes—that has occurred [7]. It has become common for one to hear someone who has been mugged say that they didn't file a police report because they thought it wouldn't bring about any effect. The idea behind WikiCrimes is to provide a common area of interaction among people so that they can make the reports and monitor the locations where crimes are occurring. It is based on the principle that the ones who hold information about crimes are the citizens. Thus, individual participation, in a collaborative manner, can generate collective intelligence. In other words, if there is active participation, crime mapping will start being done collaboratively, and everyone will benefit from having access to information about where crimes occur.

Main Architecture

The open and participative characteristics identified in WikiCrimes makes it susceptible to abuses or attempts at fraud. It is important that as many people as possible collaborate with the system, contributing to the growth of its data records, but it is equally important that the information registered in the system be reliable, so that the system can become a trustworthy source of information. By saying that, our claim is that an open collaborative system must be viewed as a kind of open multi-agent system, where a number of human and/or artificial agents interoperate pursuing their individual or common goals [12]. These agents can contribute positively or negatively to the organization and goals of the system as a whole, and the agents are free to join and leave the system as they wish, as long as they obey certain rules that must be observed concerning the management of the system.

The management of interoperation among agents is a complex task, and robust techniques and methodologies for the development of reliable and open Multi-Agent Systems (MAS) have been studied in academia [1, 2, 3, 6, 9, 11]. Such techniques are aimed at the modeling and implementation of features that give openness to those agents, allowing them to have the ultimate choice of obeying regulations or dealing with possible sanctions imposed by the MAS norms. After all, agents are

autonomous entities and the biggest challenge is to have a coordination system where the agents can be free to decide what to do but at the same time be encouraged or seduced to obey the regulations of the society they are entering.

WikiCrimes as a Multi-Agent System

We have modeled WikiCrimes as an Open Collaborative System [12]. Basically we envision the system as an Institution of Agents, as defined in [11]. In [12], some colleagues and I define a set of agents that are in charge of system tasks, a set of agents contextualized to the WikiCrimes system domain, and a set of ontologies. We describe now the most relevant agents in relation to the scope of this paper, and the ontologies will be defined in the next section. The Reputation Agent is responsible for keeping updated a database of reputation owned by the agents registered to the system. Reputation is calculated as a combination of the trust the WikiCrimes system has in some members of the institution and the behavior of the agents in relation to interactions and acceptance of the system norms. The other types of agents characterize the possible users of the system: the Registered User Agent, representing a typical user of WikiCrimes, who is able to register, confirm and disconfirm crimes, browse the environment, denounce abuse, and indicate other agents to confirm crimes; the Invited User Agent is the agent that is indicated to confirm a crime; the Certifier Entity Agent is a special kind of agent (typically representing an organization) that holds a respected position in the community. The Browser User Agent represents the users that only browse the institutional environment, basically seeking information in the system.

Interacting with WikiCrimes

The WikiCrimes interface is basically divided into spaces for map manipulation and search functions. Figure 1 depicts the main screen of WikiCrimes, which offers a crime search function (right side of the figure) that enables users to view the reports of crimes, filtered by crime type, date range and time range. The search can also be done by means of a direct search on the map by querying and viewing a location on the map. To post a criminal occurrence, the user must be registered and logged in to the system. The system currently permits the user to post the following types of crime:

- Robbery: robbery against the person, robbery against property, attempted robbery against the person and attempted robbery against property;
- Theft: theft against the person, theft against property, attempted theft against the person and attempted theft against property;
- Violent Crime: homicide, attempted homicide, robbery followed by murder (aggravated robbery)

involving death of the victim), brawls or fights, domestic violence, abuse of authority.

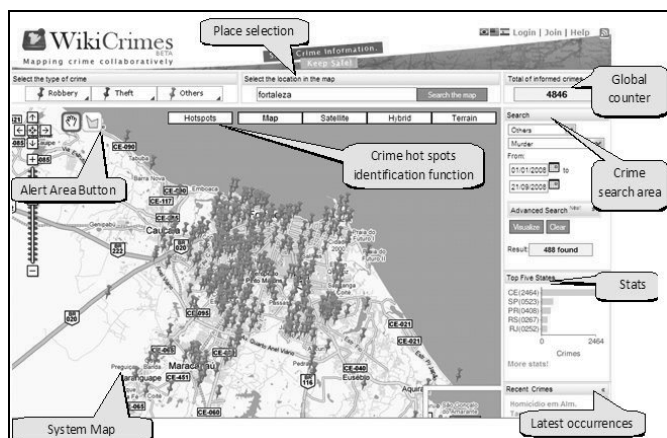


Figure 1 Overview of WikiCrimes main interface showing the map of homicides in Fortaleza with pin points and interaction areas

To obtain information about a crime, suffice it to click on the respective marker of the crime. Note that the complete description of the crime is accessible only for those who are logged in. In Figure 2, the map of crimes is also shown by means of markers. Different ways of visualization are provided, such as hot spots in which the identification of dangerous areas is based on clusters represented by the intensity of a color in a given region [4].

Credibility of the Information

In WikiCrimes there are not many requisites to become a member of the system. The only personal questions asked are a name and valid email address; no document identification is required so that people will not be afraid to post information in the system. It is up to the user to provide the system with information that increases the reliability of the crime register. It is possible to add links for videos, newspapers, photos or any other document (such as a police report) that helps another person to believe in the register. Moreover, for every criminal fact registered in the system, it is requested that there be an indication of at least one person who can confirm that the information posted is true. The more the information is confirmed, the more it is considered trustworthy. These indications for confirmation of information generate a graph where the vertices represent users of WikiCrimes and the edges represent the indication of others to confirm the criminal fact registered. The graph represents a Social Network formed by the WikiCrimes users. By doing so, we are able to create a reputation model that plays a strong role in the matter of identifying “bad agents” in the institutional environment, and the social net of WikiCrimes users is a fundamental tool for that. By “reputation,” we

mean a score that represents the view of a community about a member of that community. Trust is a score that reflects one entity’s subjective view of another. Trust can be calculated upon the acknowledgement of a given reputation.

In WikiCrimes, reputation is built upon interaction of the agents, by the observation of breaking the commitments, and consequently the norms of the system. We have thus defined the concept of reputation for the source of information, i.e. whoever registers the criminal fact. In other words, in WikiCrimes we have a social network formed by the users that are registered in the system—forming the Social Network Layer—and the information that is posted in the system—the criminal facts. The information posted in the system forms the information layer. The goal is to build a function that will calculate the reputation of the users and reflect that reputation in the trustworthiness of the information that is posted in the Crimes Layer. Some entities, such as the press and governmental agencies, are labeled as Certifier Entities, and are therefore considered very well reputed. But that is not enough; the open characteristic of the system, in the sense that anyone can be a user, does not facilitate the task of knowing the reputation of all the users of the system. The attribution of reputation to the users who are not qualified as Certifier Entities is fundamental. We can assume that an agent delegates a commitment to another agent when the former indicates the latter to confirm a crime. The act of acceptance of the commitment by the agent indicates a relationship of trust between the agents. If the commitment is broken by the second agent, it will be penalized in reputation points. This relationship suggests the adoption of mechanisms to propagate trust in the social network built in WikiCrimes. The Administrator and Certifier Entity roles have a very good reputation to start with. These agents serve as a starting point for the propagation of trust to the agents indicated by them, then to the ones indicated by those indicated, and so on, similar to the propagation of trust and distrust for demotion of web spam described in [18]. Besides the reputation computed from propagation, the reputation and trust will be updated based on the interactions identified in the MAS, namely posting a crime, confirming a crime positively, confirming a crime negatively and denouncing abuse. We named this “Acquired Reputation.” Agent interaction with the Information Layer will help to identify the agents that are breaking these social norms. These interactions will indicate how much trust the users of the system have in such information.

Ontology Representation

Semantics in WikiCrimes is represented by means of ontologies. Two basic ontologies are used by most WikiCrimes agents for representing the concepts of crime and reports of crime. A crime has a type, a time (imported

from the time ontology of [17]), an address, a type of weapon used, etc. We defined a crime ontology inspired in the Criminal Act Ontology [13] in the context of the OpenCyC Project [13]. The report of crime has information about a particular crime and about the reporting itself, such as the type of the denouncer (victim, witness or just someone who was informed about it) and whether this denouncer knows whether the fact was or was not communicated to the police. Moreover, in the report there is a set of features to describe the provenance of the information, such as the type of source and how reliable it is. For representing these latter features, we imported the PML2 ontology [8]. However, we specialized some of the concepts to cope with the particular features of WikiCrimes, particularly for representing the notion of reputation inspired in [10, 14].

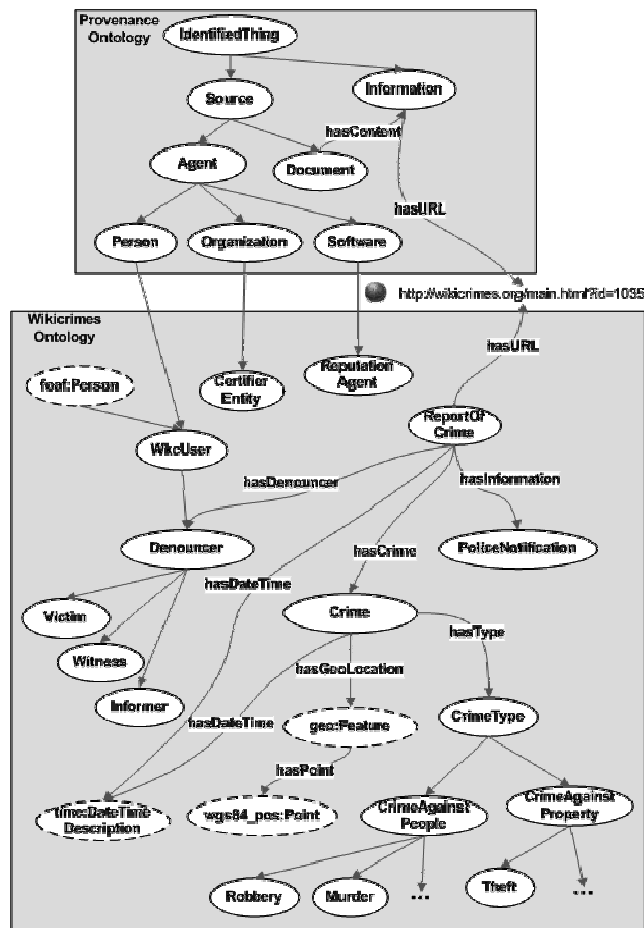


Figure 2 Crime and Report Crime ontologies and their relationship with PML and GeoNames

Figures 2 and 3 depict the classes we have defined in WikiCrimes. In order to represent provenance of information using PML, real-world entities are defined as instances of the two PML subclasses: *Information* and *Source*. The *Information* class enables the representation of the report of crime. It has properties to inform the content

of the report, its language, format, the denouncer, the date and time of the report, etc. The *Source* class allows the representation of an *Agent* or a *Document*. The *Agent* class can represent an *Organization*, a *Person*, or *Software*. An agent can be represented as a member of an organization that, in turn, has a group of members. The *Organization* class can be specialized to represent a *Certifier Entity* as well.

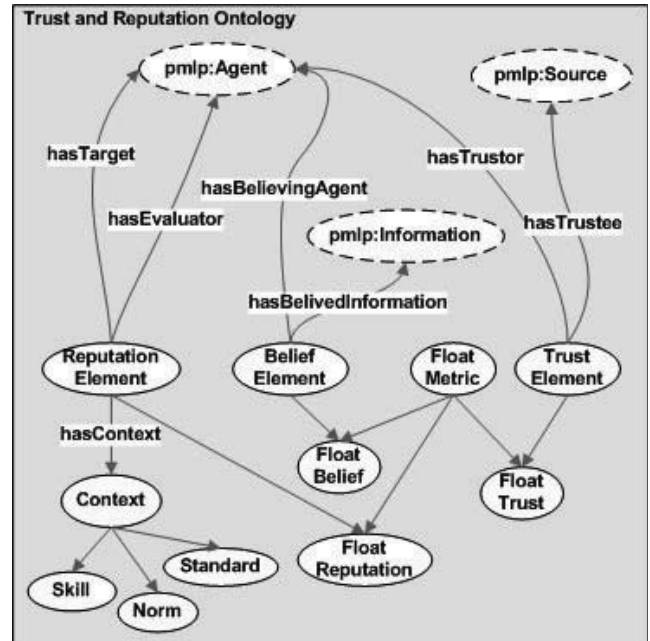


Figure 3 Reputation and Trust ontologies used in WikiCrimes

For representing the concepts related to the reputation model, we defined the concept of *ReputationElement* that includes the necessary elements to represent the spreading of an evaluation (a report on reputation). For example, "it was told that the reputation of Agent X in WikiCrimes is 0.81". Therefore the class *ReputationElement* has, as elements, the evaluation itself, the teller and receiver of the evaluation, the context and the strength of the evaluation. All the sources of the system (represented by the *Source* class) are susceptible of being evaluated or having an active part in the generation or diffusion of evaluations. In WikiCrimes, the *Context* class represents the norms that regulate the system. The *FloatReputation* class represents a value that expresses how good or bad the target agent is in a given context. The value of the reputation in WikiCrimes is represented by a real number belonging to the interval [-1,1], -1 being completely bad and 1 being completely good. The *Strength*, represented by a real number belonging to the interval [0,1], is a subjective measure set by the *Source* that indicates how reliable the evaluation is, 1 being the maximum reliability. This is an attempt to make the evaluations more accurate by the inclusion of the element of reliability in the calculations made by the *Source*. In WikiCrimes, the *Strength* is identified by the

reputation of agents, who are confirming crimes and denouncing abuse, and this influences the calculation of the Target's Acquired Reputation.

The notion of belief in information is also represented in WikiCrimes. Agents confirm a register of crime because either they trust in the agent who has sent the invitation or they believe in the information posted. Note that it is possible to confirm/disconfirm a register of crime without having received an invitation. In this case, the agent's action is exclusively based on the belief/disbelief in the information posted.

Another important concept being represented in WikiCrimes is that of location. Crimes occur in a certain geographic location, which links the crime event with several other events. For instance, the interest of a tourist in a particular city spot can be associated with security information about that place. The representation of the crimes' geographical location uses the GeoNames Ontology [4] that makes possible to add a geospatial semantic description about locations. It has over 6.2 million geoname toponyms such as cities, towns, countries, and the relation between them. The ontology includes features such as: contains (administrative divisions), neighbors (for features with a boundary) or nearby features. The GeoNames Ontology also represents latitude and longitude values in RDF format through the use of the Basic Geo (WGS84 lat/long) vocabulary [4]. In WikiCrimes, the *Crime* class has an object property called *hasGeoLocation* with *geo:Feature* as a range value. *geo:Feature* is a class of the GeoNames Ontology with properties such as country, postal code and alternate name. Finally, we mention the way we have represented the members of the social network formed by WikiCrimes users. We have imported the concepts of person and friends as they are described in the FOAF (Friend-of-a-Friend) ontology [15], the email address being the unique identifier in this case.

Putting Semantics to Work

The *Location* ontology in WikiCrimes is intended to be the manner to interconnect people from other systems that use the same concepts. People who are looking on a web site for a location on where to spend their holiday can exchange information with other sites using the same ontology, and interconnect their users offering services of the other site. WikiCrimes can offer the safest places to go or the neighborhoods to avoid, or—even better—a user from WikiCrimes can give information about some street or restaurant that he has visited, or any other target point, therefore connecting these users from different groups. Thus, WikiCrimes and its users will become potential security advisers for other web users. It is important to note that the concepts of reputation and trust are fundamental to the acceptability of a service of this type.

Another way to increase people's connectivity is exploring the formation of groups or communities of interest. In WikiCrimes, such a notion is present in two different manners. There are groups formed by the explicit connection generated when a user indicates other users to confirm a crime (supposedly a friend), and the one formed by users with common interests related to a specific geographic location or target point, such as: a street address, a school, a shop, a hotel, etc. This affinity with others based on specific locations can be captured because, in WikiCrimes, the user can demarcate an alert area in the system. Such an area is then monitored by the system and every register of crime made inside the area is informed by email to the user. Once created, this alert zone is tagged and can be shared with other registered users of the system or invited users, including users who are members of other social networks. This feature aims to increase the number of registered users of the system and expand the connection among the users, allowing for a bigger network effect.

Another context of using WikiCrimes by exploring its semantic representation can be envisioned in terms of mobile devices. The identification of the context of use (with respect to the geographic environment) where the user is inserted allows the generation of contextual alerts. For instance, an SMS message can be sent to the user's cell phone, alerting about the past occurrences of crimes in that particular drugstore. Also, the computing of safe routes can be produced on cell phones as well as GPS navigators.

Conclusion

WikiCrimes was born in Fortaleza, a city of 2.5 million inhabitants in the Northeast of Brazil. Therefore this is the place that has the large majority of users and registers of crime. More than half of the total number of crimes reported in WikiCrimes has been pinpointed there (about 2,600). Articulation with organized society, workshops, lectures and a diverse and consistent campaign of local advertising have been particularly intense in the city. The expansion to other regions is gradually occurring, mainly by means of agreements with city halls of medium-sized municipalities and with collaborators who maintain blogs on a similar theme. In addition to the multifaceted technological component that characterizes WikiCrimes, other scientific challenges are imposed and deserve special attention. Three of them are currently objects of study by the WikiCrimes team and relate to the semantic web issue. First, it is worth mentioning the studies aimed at making WikiCrimes available on mobile devices. The goal of such actions is to bring WikiCrimes closer to its users, leading them to join the WikiCrimes community at different times and to expand on mere access to the site. A prototype for operating as a mobile version to run on Nokia's S60 platform of cellular telephones is already being tested. The

second topic of investigative research is in regard to the development of software capable of reading pages of online newspapers, describing news stories about crimes and registering such crimes automatically in WikiCrimes. This involves investigating semantic methods of exploring web content as well as methods of processing natural languages. Finally, but no less important, we are making efforts to develop an algorithm for rendering the geographic information around the place of crimes by considering the semantics of the region and surrounding areas. By doing so, we can prepare a dataset to be mined regarding the patterns that can explain the reasons the crime occurred. With these three elaborations, we will also be able to extend the notion of explanation beyond the notion of provenance and trust, to take into account the reasons crimes are occurring as well.

References

- [1] Aldewereld, H., Dignum, F., Garcia-Camino, A., Noriega, P., Rodriguez-Aguilar, J.A., Sierra, C.: Operationalisation of norms for usage in electronic institutions. In: AAMAS '06: Proceedings of the fifth international joint conference on autonomous agents and multiagent systems, New York, NY, USA, ACM Press (2006) 223–225.
- [2] Arcos, J., Esteva, M., Noriega, P., Rodriguez-Aguilar, J., Sierra, C.: Engineering open environments with electronic institutions. *Engineering Applications of Artificial Intelligence* 18(2) (2005) 191–204.
- [3] Colombetti, M., Fornara, N., Verdicchio, M.: The role of institutions in multiagent systems. *Proceedings of the Workshop on Knowledge based and reasoning agents, VIII AIIA* (2002) 67–75.
- [4] Furtado, V. Ayres, L., de Oliveira, M., Vasconcelos, E., Caminha, C., Freitas, E.: Collective Intelligence in Law Enforcement – The WikiCrimes System. Submitted to *Information Science*.
- [5] GeoNames Team. The GeoNames Geographical Database Ontology. <http://www.geonames.org/ontology/>.
- [6] Grossi, D., Aldewereld, H., Vázquez-Salceda, J., and Dignum, F., Ontological aspects of the implementation of norms in agent-based electronic institutions. *Computational & Mathematical Organization Theory*, 12(2), 251-275.
- [7] Kahn, Tulio. Boletim de Ocorrência - prover para poder prever. Retrieved September, 2008 from Forum Brasileiro de Segurança Pública. Web site: <http://www.forumseguranca.org.br/artigos/>.
- [8] McGuinness, D., Ding, L., Pinheiro da Silva, P., Chang, C.: PML2: A Modular Explanation Interlingua. In *Proceedings of the AAAI 2007 Workshop on Explanation-Aware Computing*, Vancouver, British Columbia, Canada, July 22-23, 2007.
- [9] Minsky, N., Ungureanu, V.: Law-governed interaction: a coordination and control mechanism for heterogeneous distributed systems. *ACM Transactions on Software Engineering and Methodology (TOSEM)* 9(3) (2000) 273–305.
- [10] Mollenkopf, J. Goldsmith, V. McGuire, P. Sara, McLafferty.(2000) Identification, development and implementation of innovative crime mapping techniques and spatial analysis. Washington, D.C.: U.S. Department of Justice, p. 27.
- [11] De Oliveira, M., Purvis, M.: Aspects of Openness in Multi-Agent Systems: Coordinating the Autonomy in Agent Societies. In *Intelligent Integration in Distributed Knowledge Management*. Ed.: Krol, D. and Nguyen N. T. New York (2008) 116-128.
- [12] De Oliveira, M., Furtado, V., Cranefield, S. and Purvis, M. Open Collaborative Systems as Institutions of Agents. In *Proceedings of IEEE/WIC/ACM International Conference on Intelligent Agent Technology (IAT 2008)*. IEEE Press, (2008).
- [13] OpenCyc Team. OpenCyc Knowledge Base. <http://www.opencyc.org/>.
- [14] Pinyol, I., Sabater-Mir, J., Cuni, G.(2007) How to talk about reputation using a common ontology: From definition to implementation. In: *Proceedings of the Ninth Workshop on Trust in Agent Societies.. Hawaii, USA*. pp: 90-101.
- [15] The FOAF Project Team. <http://www.foaf-project.org/>
- [16] W3C Semantic Web Interest Group. Basic Geo (WGS 84 lat/long) Vocabulary. <http://www.w3.org/2003/01/geo/>.
- [17] W3C Semantic Web Interest Group. Time Ontology in OWL. <http://www.w3.org/TR/owl-time>.
- [18] Wu, B., Goel, V. and Davison B. D., Propagating Trust and Distrust to Demote Web Spam. In the proceeding of *Models of Trust for the Web Workshop (MTW)* (2006)