WeKnowIt Annual Report 2009

www.weknowit.eu

European Commission
Information Society and Media

Digital content generation and sharing between users and on-line applications is rapidly reaching a mass that makes relevant information extremely complex and costly to handle. Intelligent processing and management of such information can enable efficient access and exploitation of the underlying knowledge. The main objective of WeKnowIt is to develop novel techniques for exploiting multiple layers of intelligence from user-contributed content, which together constitute Collective Intelligence, a form of intelligence that emerges from the collaboration and competition among many individuals, and that seemingly has a mind of its own. To this end, input from various sources is analysed and combined: from digital content items and contextual information (Media Intelligence), massive user feedback (Mass Intelligence), and users social interaction (Social Intelligence) so as to benefit end-users (Personal Intelligence) and organisations (Organisational Intelligence).

Summary of Activities

The project is at the end of its first iteration development phase, and R&D has focused on integrating the first prototype of WeKnowIt for the two use cases: an *Emergency Response* application, and a *Consumers' Social Group* application. The infrastructure for the platform is complete and state of the art components from the various Intelligence layers are finalised, adapted to the platform requirements and fully integrated. As a result, the first prototype is up and running with the initial versions of the two use cases been built upon it. All procedures on testing and evaluation have been defined to guarantee smooth transition to the prototypes testing and evaluation phase. In parallel, a common understanding has been established in the consortium for the methodological approach of Collective Intelligence within WeKnowIt, and the way that it will be provided. During this year the initial exploitation plan was completed, and the project was involved in various dissemination and clustering activities. In the next year, the project will enter the 2nd implementation cycle, continuing with integration of more and improved Collective Intelligence approaches within the prototypes, and also by combining components from different Intelligence Layers.

User requirements and WeKnowIt relevance

The Collective Intelligence approach in WeKnowIt is applied into two distinct demonstrable applications:

- an *Emergency Response* application, where user-provided intelligence about large scale emergencies can be leveraged into informed knowledge for the efficient planning of emergency-response actions, and

- a *Consumers' Social Group* application, where consuming activities of end-users are exploited for the extraction of information for the organisation of similar personal or group activities (e.g. organisation of travel events).

WeKnowIt, with respect to its implementation technology and provided applications, is expected to target end-users that combine any of the traits, as listed in the table below.

User traits	User Needs/Requirements	WeKnowIt relevance
Technology Literacy	High demands on sophisticated services and on the quality of their experience	Provide an enhanced user experience through integration of latest Web 2.0 and Collective Intelligence mass content analysis approaches
Environmental Awareness	The wish to contribute against devastating events, as a result of environmental sensitivity	The implementation of an ER management application based on user-contributed information
Modern Traveller	A global increase of traveller population, that plan and manage their trip activities through internet sites	The implementation of a traveller-guide application for planning and organising the trip activities both beforehand and during the trip including content analysis techniques
Network Agnostic	The need for un-interrupted and user-friendly services, unrelated of any prerequisite for underlying technical knowledge	The unremitting servicing of end-users through the seamless deployment of web technologies and data access bearers
Socialising	People's natural tendency of socialising, meeting and interacting with new people, or forming virtual worlds and networks of users with similar interests	The incorporation of personalised social networking features in the implementation of provided applications
Multi-interactivity	The management requirement of large data volumes, deriving from diverse, yet complementary sources of information	The semantic-based incorporation of many layers of intelligence act as an *autonomic* organise mediator between the user and the world
Trust consciousness	User reassurance need about the secure handling and protection of sensitive and personal data	The implementation of secure shells for transmission of sensitive data and the user's complete control over the personal information that becomes shareable
Ubiquitous service	The anytime/anywhere delivery of services	The provision for fixed as well as mobile end-users while exploiting the ubiquity of beyond-3G networks. Further, the implementation of a ubiquitous computing environment will promote the provision of a context-awareness pervasive system
Access media convergence	A unified (i.e. non-fragmented) means of service access, despite the multiplicity of end-terminal divergence	The design and development of a converged and user-agreeable access interfacing to the end-users

From various layers of Intelligence towards Collective Intelligence

The WeKnowIt approach towards Collective Intelligence builds on two aspects: mass content availability provided by a lot of users and availability of analysis techniques and results from different layers. These are combined at different levels (content, results) and with various techniques (sequential, late fusion) in order to produce additional, higher-level information, more reliable by eliminating false alarms.

Part of WeKnowIt is exploring how *Personal Intelligence* can contribute to Collective Intelligence by providing and accessing of information by individuals who have different informational needs and access requirements. A graphical login mechanism has been developed which allows mobile users of WeKnowIt technologies a simple means of accessing protected services. WeKnowIt has also built a web display framework, Sparks, which supports the flexible display of information by determining how to display it on the basis of the information properties.

Innovative features within the *Media Intelligence* field include intelligent, automated content analysis and retrieval techniques, based on visual image similarity techniques, as well as advanced speech and textual content processing. Significant efforts have been spent also on visual image retrieval and localization of digital still images (Fig. 1), facilitating efficient features and keypoints extraction (Fig. 2) and matching algorithms for classification to potentially identifiable objects, such as well-known landmarks. Information from external sources (e.g. Wikipedia links of suggested tags) is also exploited in the process (Fig. 3).

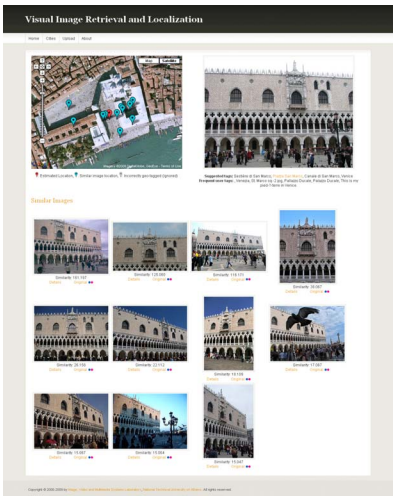

Fig. 1: Retrieval of similar images and localization of query image on Google Maps.

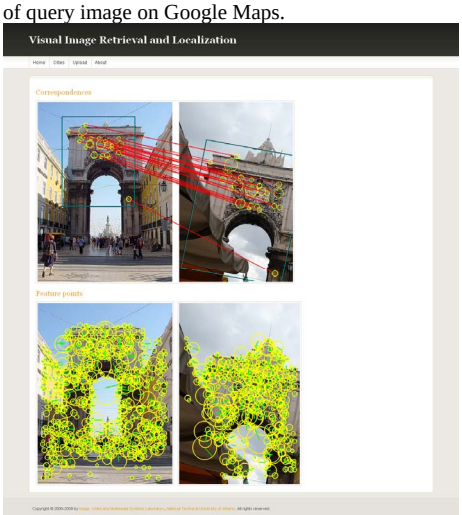

Fig. 2: Key-point extraction under almost any viewpoint and geometrical transformation.

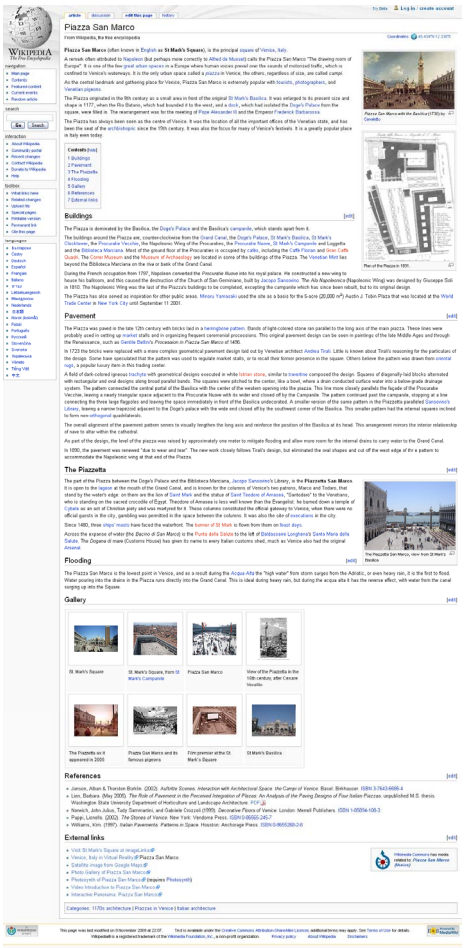

Fig. 3: Utilization of external web sources, such as Wikipedia, to enhance recommended tagging information for the query image.

Advancements in *Mass Intelligence* include extraction of hidden features from massive data, utilization of information stored in tags and social networks, modelling dynamics of items associations' using user feedback, and practical faceted navigation in the large scale knowledge base. Significant effort has been spent on discovery and analysis of latent topics in document corpora (Fig. 4), and on identifying communities of related tags in large folksonomies (Fig. 5). Another approach to improve search experience in large-scale dataset is discovering facets and ranking them based on relevance to the user (Fig.6).

Britney Jean Spears (born December 2, 1981) is a Grammy Award-winning American pop singer and entertainer. Raised in Kentwood, Louisiana, Spears first appeared on national television in 1992 as a contestant on the Star Search program, and went on to star in Disney Channel's television series The New Mickey Mouse Club from 1993 to 1994. In 1997, Spears signed a recording contract with Jive, releasing her debut album …Baby One More Time in 1999 […]

Latent Topics			
204	102	1201	47
70%	32%	9%	2%
music	female	television	international
artist	actress	radio	national
pop	movie	noise	country
singer	award	commedy	birth
album	club	horror	men
record	singer	channel	woman

Experts
lucas86
anna_rio
musicman
miro_1986
abbabbac
johnston
ziggy_zaggy

Fig. 4: Latent topics and expert users identified for given text.

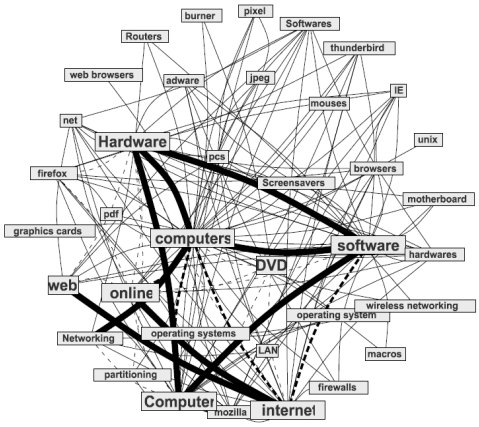

Fig. 5: Tag community around 'computers'.

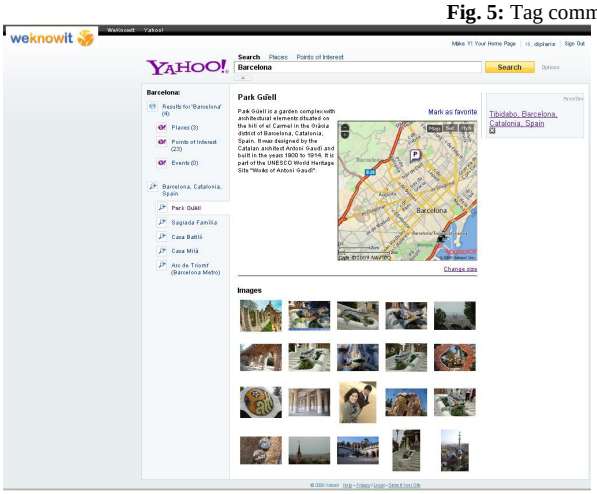

Fig. 6: Faceted navigation specialized for places integrated by Yahoo! with WeKnowIt Consumer Social Group first prototype.

Almost any application, online or offline uses authorization. Who is allowed to perform which tasks? Who can access which file? The Authorization Design Language (ADL) developed as a WeKnowIt *Social Intelligence* component is a formal language created to face this issue. Its flexible design allows representing various conditions under which access shall be granted. May a user access a resource only from the intranet? Can a file be accessed "read only" via a handheld? The ADL is capable of defining such conditions. It can be used as a general authorization database for defining, maintaining and querying authorizations. By this, we expect the ADL to operate as a middleware between applications' business logic and their authorization policy. WeKnowIt also developed a software toolbox for the analysis of social networks. Exemplary questions that can be tackled with the help of this software are: What is the structure of this network? Which people form a close group? Which people provide, moderate or consume information? The application focus is on large, evolving networks like we deal with in emergency situations or active on-line communities.

For the *Organisational Intelligence*, advancements have been achieved with developing the pattern-based ontology for events, the core ontology Event-Model-F. The Event-Model-F provides a formal representation of the different aspects of events in which humans participate such as time and space, composition, correlation, and documentation. Compared to existing

models, the Event-Model-F differs in providing sophisticated support for modelling causality, correlation, and interpretation of events (Fig. 7). The Event-Model-F is designed to be domain independent, which means that it can be applied for emergency response (Fig. 8) as well as tourism but also for other domains. It is employed in the SemaPlorer application for the faceted navigation in a very large, distributed and heterogeneous semantic data set (Fig. 9). People participating in events may be organised in different companies or communities. Here, a distributed group management has been developed that bases on the well-known Friend-of-a-Friend vocabulary. Events and the activities of professional emergency response entities during an emergency incident are documented in log-files. A log-merger application processes these log-files with NLP-techniques to provide the users access to the emergency response related information like location, role, and action (Fig. 10).

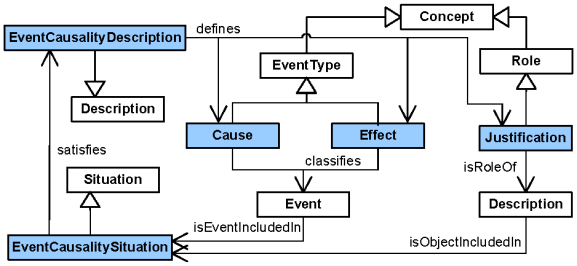

Fig. 7: Causality Pattern of the Event-Model-F

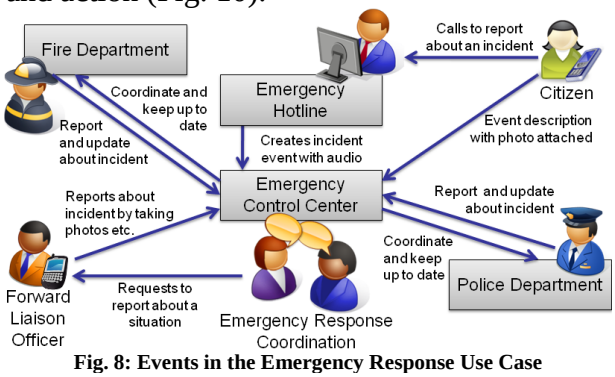

Fig. 8: Events in the Emergency Response Use Case

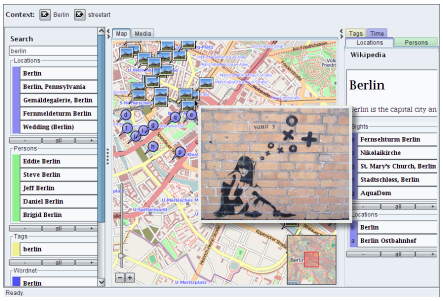

Fig. 9: SemaPlorer for the faceted navigation in a very large and semantically heterogeneous semantic data set

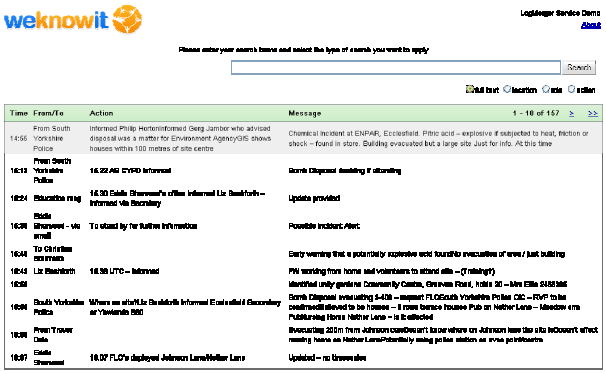

Fig. 10: Analysis of emergency response log-files and search for location, role, and action

System specifications and infrastructure technology

The software framework created for the WeKnowIt System serves as an integration basis for the modules produced by each intelligence layer. It provides a Service Oriented Architecture environment (realized using Java technology) based on the stack of popular open-source projects (Fig. 11). The WeKnowIt System provides access to the integrated data storage (Fig. 12), that combines different types of data storages (i.e. RDBMS and triple stores), and is prepared to store and effectively serve data of various kinds - multimedia files, relational data, Java objects and RDF triples. On top of the provided software framework, use case applications are built, using services provided by the modules integrated within the WeKnowIt System. The use case applications communicate with the WeKnowIt System via web services. Additional layers of abstraction of the WeKnowIt System, prepared per use-case, result in easy to use API tailored to the requirements of each use-case application.

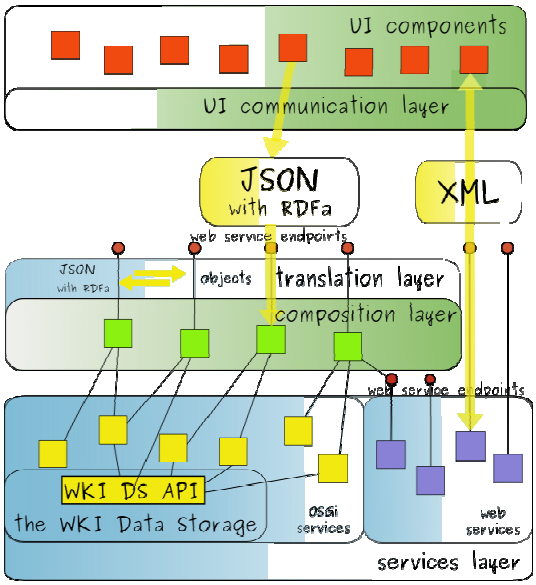

Fig. 11: The WKI System architecture - with use-case applications

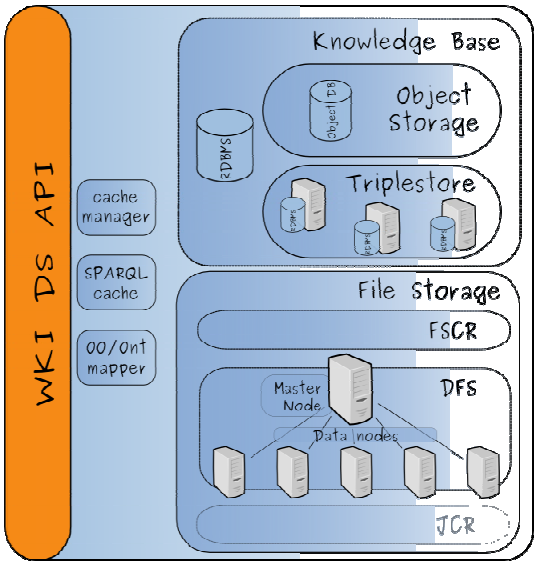

Fig. 12: The WeKnowIt Data Storage

The WeKnowIt use cases

Emergency Response case study

The purpose of this case study is to show how WeKnowIt technologies can support individuals and organizations in times of emergency. The case study examines how citizens can provide information to the Emergency Services using an intelligent upload process. The intelligent upload process ensures that the information the citizens upload (for example uploading an image of a flooding incident) is enriched through the intelligence layers present in the WKI system. The cumulative information uploaded by citizens is then used by the Emergency Services in making planning decisions for dealing with the emergency. In addition to this decision making, the prototype allows the Emergency Services to make selected information available to the general public thereby allowing the intelligence exchange of information between the citizens involved in an emergency incident and the organizations that are dealing with the emergency.

Fig. 13: User interface on mobile phone for Emergency Response case study

The intelligent upload process is supported via a Web application that features a mobile interface. This allows users to provide information, tag that information using suggestions generated by automatic normalization to domain ontologies and to get updates on the incidents to which they are providing information.

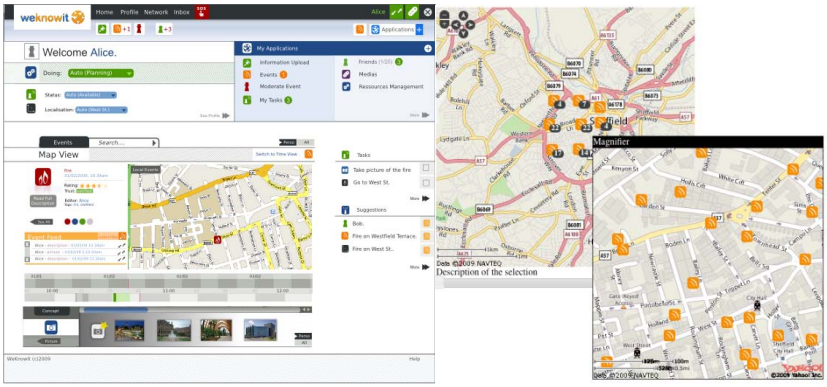

Fig. 14: User interface on computer desktop for Emergency Response case study

The emergency personnel are then able to view the uploaded information in a geographical display. The geographical display allows them to filter the information by location and by time ensuring that they have access to up-to-date and relevant information when making decisions.

Consumer Social Group case study

This case study covers a scenario where the focus is on travel or one-day cultural trip events. The scenario is composed of two parts: The first part involves preparing of the journey. In this phase the user identifies places, points of interest, activities or events. The first prototype features a rich internet application to display information related to the travel, search results based on Mass Intelligence components, localization of points over maps to enrich the user's context, display of images related to the selected query, location or point of interest, and offers three types of searches: in the entire collection, by location and by point of interest. The application also offers a favorites storage functionality.

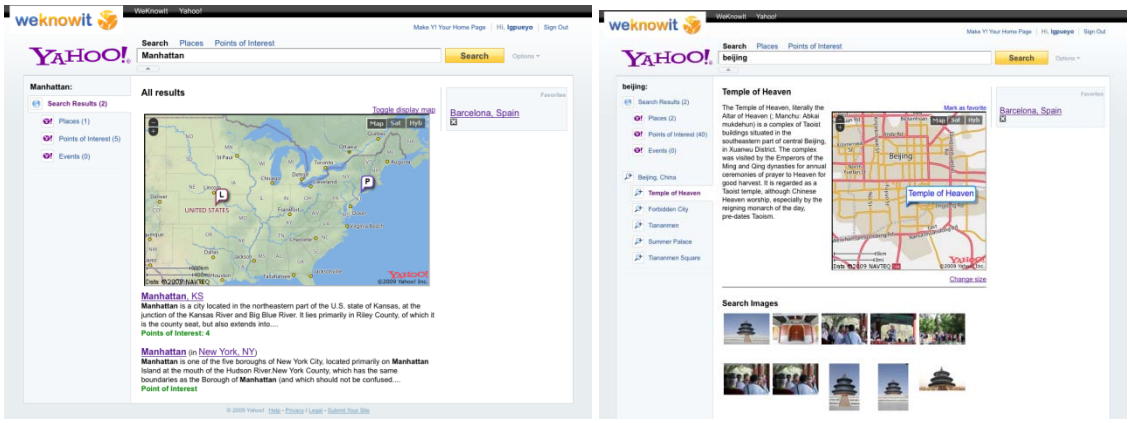

Fig. 15: User interface on computer desktop for Consumer Social Group case study

The second part of the scenario involves getting mobile guidance during the trip. In this phase, the user gets useful and helpful information on-line on the mobile phone; besides he can take pictures recording pieces of his experience. The first prototype allows for a mobile version of searching for places and points of interest, providing detailed information of the places and points of interest found, making use of mass content analysis techniques in Wikipedia articles and other public sources.

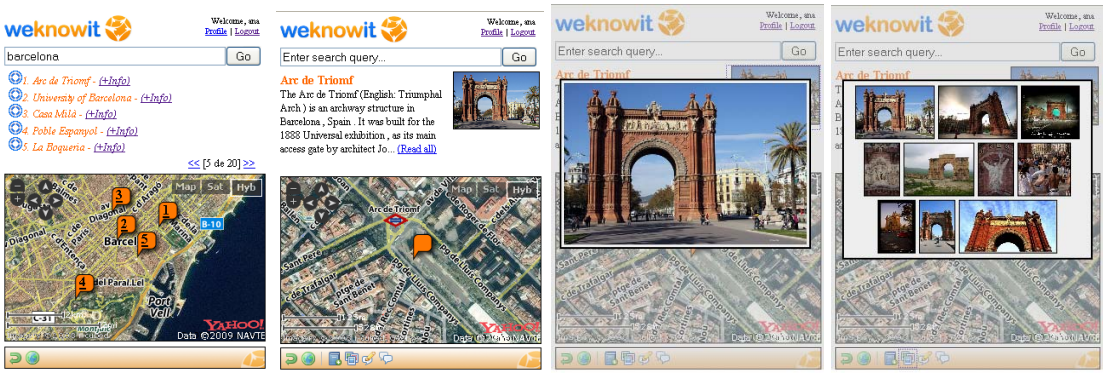

Fig. 16 : User interface on mobile phone for Consumer Social Group case study

Both case studies, being at the end of the first iteration, are aiming to evaluate the features developed with actual end users. This evaluation is taking place during December 2009 and January 2010 in Barcelona (Spain).

User Involvement, Promotion and Awareness

During 2009, WeKnowIt has participated and organised a number of events, in order to promote its progress and raise the awareness of both the scientific community and the general public about its objectives. Highlights among these have been the successful organisation of a *Summer School on Multimedia Semantics*, hosted by the University of Koblenz-Landau, the organisation of the 1st international workshop on *Collective Knowledge Capturing and Representation*, held in conjunction with the 5th International Conference on Knowledge Capture (K-CAP 2009), as well as the organisation of the 1st international workshop on Events and Multimedia, held in conjunction with the ACM multimedia 2009 conference, in Beijing, China. Also a number of publications in conferences and journals has been achieved[1]. Liaison activities have been initiated with PRONTO project, a recently started research project, funded under EU's 7th Framework Program. PRONTO shares much of WeKnowIt research interests with applications in the field of Emergency Response Management, a use case to be jointly exploited in favour of the complementary and collaborative progression of common evaluation and dissemination activities.

Future Work

The following year starts with the initiation of the 2nd cycle of WeKnowIt design, development and evaluation. The phase will include design and implementation in all Intelligence layers, but most importantly in developing new Collective Intelligence techniques, by combining existing or new approaches from the different Intelligence layers. Results from the 1st evaluation phase will serve as valuable input for defining the requirements for the final versions of the two use cases applications. Moreover, extra evaluation tests will be performed with external users of the platform, especially in terms of collaboration with other similar domain EU projects.

Further Information

The consortium

1. Centre for Research and Technology Hellas (Greece)
2. University of Koblenz-Landau (Germany)
3. The University of Sheffield (United Kingdom)
4. University of Karlsruhe (Germany)
5. Vodafone Panafon - Hellenic Telecommunications Company S.A. (Greece)

[1] A complete list of publications can be found on http://www.weknowit.eu/publications

6. Software Mind S.A. (Poland)
7. The Sheffield City Council (United Kingdom)
8. Brno University of Technology (Czech Republic)
9. Telefonica I+D (Spain)
10. Yahoo! (Spain)

Contacts and Information

For further information visit the project web site www.weknowit.eu or send email to the project coordinator Dr. Yiannis Kompatsiaris ikom@iti.gr.

You can also visit some of the applications that are part of the WeKnowIt prototype:
Semaplorer: http://btc.isweb.uni-koblenz.de/
The event model: http://isweb.uni-koblenz.de/eventmodel/
The ViRaL tool: http://www.image.ntua.gr/iva/tools/viral
The desktop application of the Consumer Social Group case study: http://weknowit.research.yahoo.com/csg