Template

PROJECT FINAL REPORT

Grant Agreement number: 215453

Project acronym: WEKNOWIT

Project title: Emerging, Collective Intelligence for personal, organizational and social use

Funding Scheme: Large Scale Integrating Project (IP)

Period covered: **from** 1st April 2008 **to** 31st March 2011

Name of the scientific representative of the project's co-ordinator[1], Title and Organisation:
Dr. Yiannis Kompatsiaris, Senior Researcher, Informatics and Telematics Institute, Centre for Research & Technology Hellas

Tel: +30 2311 257 774

Fax: +30 2310 474 128

E-mail: ikom@iti.gr

Project website address: http://www.weknowit.eu

[1] Usually the contact person of the coordinator as specified in Art. 8.1. of the Grant Agreement.

4.1 Final publishable summary report

4.1.1 Executive summary

WeKnowIt is a 3 year IP developing novel techniques for exploiting multiple layers of intelligence from user-generated content, which transform the large-scale and poorly structured Social Media to meaningful topics, entities, points of interest, social connections and events.

In this document the work performed during the WeKnowIt project is reported. The project overall research and development approach was iterative and organised in two distinct cycles, each one including research, integration, prototype building and evaluation phases. In this context research was centred around five main pillars, namely Personal, Media, Mass, Social and Organisational Intelligence.

Research results in each one of these Intelligence layers have been combined in order to produce results with higher impact than these produced by the mere accumulation of the individual intelligence layers results, thus leveraging the so-called Collective Intelligence.

Much effort has been spent in developing innovative Collective Intelligence technologies as well as in integrating them into the WeKnowIt System. Such approaches involved the cooperation of different Intelligence Layers. The produced scientific results were put into use in two case studies, one related to Emergency Response and another supporting a Consumer Social Group scenario. 7 prototypes have been developed in terms of the two case studies, while another 13 prototypes and demos were independently produced. All prototypes have been thoroughly tested and evaluated during the two implementation cycles of the project.

Dissemination has progressed successfully, with the organisation of several workshops, conferences and events. The consortium has managed to publish 12 articles in scientific journals, 65 papers in conferences and workshops, as well as two book chapters. In the exploitation level, WeKnowIt has demonstrated its achievements by participating in scientific and industrial exhibitions and events. Also 9 patents have been filed, one spin-off company was established, a follow-up grant by the UK

government was achieved, and many bilateral agreements on future synergies between partners of the consortium and external companies were arranged.

Overall the project has built on its strong scientific results, and with enhanced communication and interaction between partners, it has resulted to a strong Collective Intelligence platform able to support the two use cases, but most importantly capable of easily extending the application of the developed technologies to various other market, societal and scientific domains.

4.1.2 Project context and objectives

Exploitation of the knowledge hidden in the user contributed content needs scalable and distributed approaches able to handle the mass amount of available data and generate an optimized 'Intelligence' layer, also called Collective Intelligence. The key purpose of WeKnowIt is to develop novel techniques for exploiting multiple layers of intelligence from user-generated content, which transform the large-scale and poorly structured Social Media to meaningful topics, entities, points of interest, social connections and events. 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).

Thus, the major strategic objective of WeKnowIt may be summarized in extracting and leveraging Collective Intelligence by understanding mass user-generated content with emphasis on integration and bridging (e.g. social and content dimensions) and the mobile and organizational - business aspects.

WeKnowIt demonstrated the applicability of its achievements through two case studies: (i) an Emergency Response case study where users can provide intelligence about large emergencies of different scale, empowering a more effective and informed emergency. (ii) A Consumers Social Group case study providing enhanced publishing tools to support group activities (e.g. one day cultural trips), and the ability to extract meta-information from content sources and groups discussions to leverage Collective Intelligence for private, commercial and public purposes.

In this context we may identify the project major objectives with respect to the individual WeKnowIt Intelligence layers and components, as follows:

- The goal of the Personal Intelligence layer is to model user preferences in order to enable personalised effective and efficient interaction with the WeKnowIt applications using computers (e.g. via Internet), but also devices with limited capabilities in terms of user interaction (e.g. mobile phones, PDAs, etc.). This happens both during the phase of information uploading and gathering and for the information retrieval phase, where information is passed to the single user. In this context it is first needed to establish the state

of the art in Personal Intelligence, in order to design, develop and test methodologies and technologies for managing Personal Intelligence in a multimodal and multimedial context. Interaction models and scenarios needed to be defined to exemplify a personalised and context-based interaction between the users and the applications. In the following technologies for users provision of information enabling upload of content and access had to be developed. These technologies are necessary in order to allow user interaction with the system either as Collective Intelligence consumer (e.g. for ER personnel) or contributor (e.g. user reporting about an event). Modeling and representation of such activities is an important aspect, requiring models for User Interaction, Events, Context and Trust. Among the objectives of Personal Intelligence was also the development of semantic-aware visualization techniques for the prototypes user interfaces and services for user authentication and personalized recommendations. One further objective in the Personal Intelligence layer was the development of technologies in order to provide focused interaction with the user, by focusing on user and context modelling, as well as on dialogue management during the interaction. Moreover, relevant tools that utilise these models had to be developed in order to be usable in the use cases. Another line of research in Personal Intelligence was profile mining in order to build user profiles and a profile-based recommendation engine. In order to be useful for the CSG use case, techniques that fuse available city information with user profiles were required in order to provide personalised point-of-interest recommendations.

- Media intelligence is the intelligence originated from digital content items (images, video, audio, text) and contextual information analysis, either provided by the user or pre-existing, and their merging. For this purpose, intelligent, automated content analysis techniques are used for different media to extract knowledge from the content itself. Since the amount of data is large and noisy, machine learning, data mining and information retrieval methods are be used. Also the methods are able to fuse information from different sources/modalities, contextual information (e.g. time, location, and EXIF metadata), personal context (profile, preferences, etc.) and social context (tagging, ratings, group profiles, relevant content collections etc.). In general, the Media Intelligence objectives include spending of research and development efforts on algorithms and techniques for automatic analysis of heterogeneous user-submitted multimedia content, exploiting three available sources, namely still images, speech and text fragments. The development of media analysis tools resulted in a visual retrieval and localisation tool for user-uploaded still images, efficient speech OOV words detection and extraction and categorization of important textual events. Specifically, objectives of Media Intelligence included research and development on algorithms and techniques for automatic analysis of heterogeneous user-submitted multimedia content, exploiting three available sources, namely still images, speech and text fragments and the development of fusion services that are able to exploit content from different modalities. An important objective was also the integration with other WPs including integration of social

intelligence into media analysis tools and vice versa and new clustering models, which take into account social, media (i.e. content-based) and tag information for mass intelligence generation. A final objective of Media Intelligence was the evaluation of the developed techniques outcome, including comparison against the current state-of-the-art.

- Mass intelligence analyzes user feedback. Mass analysis enables input information clustering and ranking as well as information and event categorization. Also, bursts of information can be detected that may indicate potential events (emergency) and trend analysis and prediction. Facts and trends are recognized and modelled by interpreting user feedback on a large scale. For instance, a single road being blocked in a storm may not be very critical, but all access roads being blocked towards a hospital centre may be very critical in the case of an emergency. Objectives of Mass Intelligence included research and development of categorization algorithms that can be applied for massive datasets to extract additional, hidden information that is present or can be inferred from the massive user contributions. More specifically, the focus was on creating the prototype of mass question answering and included analysis of question-answers dataset with all its features. As a result, there were developed methods for creating networks of related tags, discovery of hidden topical groups based on latent topics, assigning topical experts or measuring quality of provided answers. Such work was motivated by scenarios from the WeKnowIt project and developed methods are implemented into services that can be integrated into the final system. In the following Mass Intelligence aimed at developing solutions including faceted navigation that uses facets ranking mechanism, text categorization, tag recommendation, points of interest (POI) recommendation, low-level spam detection, tag community detection and new meta-clustering techniques. As a final step, the Mass Intelligence layer goal was to develop techniques on named entity discovery and ontology-based categorization, event detection in folksonomies and evolution of topics over time in social media, characterization of social media by combining text features with spatial knowledge, event and landmark detection in social media content, user and tag activity analysis over time in social media environments, faceted exploration of image search results, automatic localization of Wikipedia articles, and dynamic analysis of cluster evolution for burst detection.

- Social Intelligence is the exploitation of information about the social relations between members of a community. Nearly everything humans do, they do in a social context because they communicate, collaborate or in some other way interact with other people. Information about the various types of social relations may be represented in communication networks, friendship networks or organization charts. First the Social Intelligence work package for the first period of the project had to investigate the needs of users in general and with respect to the use cases with a focus on social networks and the information that could be generated out of this network structure. The one objective was to create a social network analysis toolbox to

provide the service to other work packages thus contributing to collective intelligence. This was achieved by first implementing an initial analysis tool that provides basic indices like degree centrality and betweenness centrality as well as an improved clustering algorithm to detect cohesive subgroups based on the linkages in the network. The other objective was to start research into and implementation of a community administration platform that allows users in both use cases to create communities as their need may be. A further goal was to conduct research on the topics of authentication, user interface, user modelling (in collaboration with Personal Intelligence) and with Media Intelligence for including results of social analysis into media analysis. A further objective of Social Intelligence was the development of services, which can be used in the project scenarios and prototypes and exploit the results of Social Intelligence. A final goal in Social Intelligence was to explore the possibilities of the cross-usage of intelligence for the improvement of social and community services. In this context, collaboration with other WPs on the topics of user rating (Personal Intelligence), social media analysis (Media Intelligence), massive clustering and recommendation techniques (Mass Intelligence) and user group management (Organisational Intelligence) were required in order to develop cross-intelligence techniques able to recommend content items and users, or assign reputation scores to them. A further objective of Social Intelligence was the development of a social networks visualization and navigation community browser.

- Organisational Intelligence allows support of decision making through workflows exploiting the generated knowledge and taking into account existing procedures within an organisation. This is quite a departure from traditional methods where knowledge is produced by the individual knowledge worker and collected and integrated manually in knowledge based systems or organisational repositories. Objectives of Organisational Intelligence were first to enable efficient access to the extracted knowledge for the support of decisions and services within organisations. In order to explicitly represent the extracted knowledge, i.e., the Collective Intelligence in terms of events and objects in the WeKnowIt project and to support the complexity of the two use cases on emergency response and consumer social group a common event model had to be created. The model provides support for explicitly representing participation of persons and objects in events, composition of events, causal and correlation relationships, documentation of events, and most notably different interpretations of events. By this explicit representation, we allow for an efficient access to the Collective Intelligence extracted. In the following a knowledge management methodology for the WeKnowIt project had to be defined, as well as the development of a knowledge sharing interaction methodology. To develop policies for a networked definition of organisations and communities, the formal model of events (Event-Model-F) based on the foundational ontology DOLCE and also a Core Ontology for Multimedia Metadata was developed. To develop the modelling of core ontologies and aligning core ontologies for the use in

organizational intelligence, a method for the integration of existing metadata formats and metadata standards into the Multimedia Metadata Ontology (M3O) had to be implemented. A relevant demonstrator, namely the STEVIE mobile tool, was developed. Another objective in the organisational intelligence layer was the development of the user interaction ontology towards the final system prototype, aimed at preserving organisational intelligence about incidents by allowing users to post information about incidents from mobile and desktop devices. A final objective was the evaluation of knowledge sharing methodology through the evaluation of a relevant collaborative post incident management tool that has also been developed.

- In order to leverage Collective Intelligence a new methodological approach able to combine the different intelligent layers, and exploit their interactions and synergies in order to effectively harness Collective Intelligence in the integration level had to be developed, integrating the results of work from the different services into a cohesive system. The objectives of the project for Architecture & Integration included first a preparation of a system platform for integration of all the tools and components developed in project (preparation of guidelines, technical infrastructure and integration plan of future development activities is also a part of this task). A second goal was to develop the WeKnowIt Data Storage for storing of all data of the system – both the semantic data and multimedia files. A third task involved gathering of data required for the project. Then the architecture infrastructure had to be finalised and the integration of all research workpackages services for the implementation of the two prototypes in collaboration with the use cases workpackage. This task involved continuous and support to the developers. In addition, the evaluation of the WKI Data Storage, a component able of storing of all data of the system – both the semantic data and multimedia files, was another objective. A final Integration objective was the publication of guidelines for developing services according to the WKI architecture.

Regarding the two Use Cases, the main objectives were the implementation (in two cycles) of the two use case prototypes based on the available services integrated in the WeKnowIt architecture together with the necessary GUI, hardware (e.g. mobile phones), content and rest of infrastructure. Another very important objective was the technical and user evaluation of the prototypes.

With respect to dissemination and exploitation activities the project aimed at planning a summer school, publications and organisation of events. Specific objectives towards exploitations included activities such as participation in industrial events and contacts in order to establish synergies. Another objective was dissemination activities through publications in well-known conferences and journals, participation and organisation of scientific events, and through participation in industrial events and exhibitions. Also objectives towards exploitation included further activities such as contacts in order to establish synergies, patent applications, and establishing of spin-off companies, among others.

4.1.3 Main S&T results

The implementation of the project first involved the collection of the initial requirements. This phase led to the architecture specification and the application scenarios definition. The latter guided the subsequent development phase in all Intelligence Layers. In this manner, a basic set of architecture services and technologies were defined in order to start producing the prototypes. In the following, research activities were carried out within the various Intelligence Layers towards the implementation of the various aspects of Collective Intelligence. Based on this research, started the development of the services necessary for the implementation of the WeKnowIt prototypes for the two use cases: the Emergency Response application, and the Consumers' Social Group application. This was possible, after the finalization of the service-oriented architecture and integration infrastructure. An important milestone was the implementation of the first prototypes with the necessary User Interfaces and content and their technical and user evaluation. Then the second cycle of research, integration, and evaluation was initiated. Research was directed towards implementing more Collective Intelligence techniques, while a continuous integration plan was carried out for the development of the two use cases prototypes. Existing Collective Intelligence technologies were integrated into the WeKnowIt system and in the two prototypes, where new functionality features were enabled. The initial versions of the two prototypes were successfully field tested during the first evaluation cycle, a procedure that has generated more user requirements for the prototypes final versions. The integration of new Collective Intelligence services and the new user requirements have eventually resulted to the finalisation of the two uses cases prototypes. The Emergency Response enables individuals to upload information to the WeKnowIt system about an emergency incident using mobile or desktop devices, allowing intelligent enrichment of the information. This information is then presented to citizens and ER personnel allowing them to understand the incident and make improved real-time decisions and post-incident analysis on the basis of this information. The Consumer Social Group prototype supports travellers in one-day cultural trips by providing desktop and mobile applications in four stages, i.e. an online application during the travel preparation phase, a mobile guidance application during the trip, a landmark recognizer mobile tool used when the traveller is on the go, and finally an online post-travel photo management application. All procedures for testing and final evaluations were efficiently scheduled and the final evaluation phase of the two prototypes was organized and performed. The Emergency Response prototype was evaluated in three rounds: an emergency simulation workshop, a longitudinal user study and citizens evaluation. The evaluations took place in Sheffield and Krakow. The prototype was evaluated by 37 users in total. Concerning the Consumer Social Group case study, the Travel preparation tool has been evaluated in Athens and Krakow from 13 evaluators testing the prototype and answering relevant questionnaires. The Mobile guidance application has been evaluated in Madrid, Krakow and Athens from 13 evaluators, who have participated in 4 hour field tests in the streets of these cities. The WKI image recognizer tool has been evaluated in Krakow, where evaluators tested the efficiency of the tool in recognizing landmarks of Krakow. In general, the feedback collected about

the concepts implemented for both case studies in research prototypes is positive, though areas of improvement have been identified.

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.

An overview of the results according to various aspects and dimensions of WeKnowIt and Collective Intelligence can be seen in Figure 1. The various services developed by each Intelligence Layer, which are presented in detail in Section 3 are shown categorized as:

- **Processing** being research and services applied for content analysis and structuring generating information useful for the WeKnowIt use cases:
    - Locations
    - Topics
    - Entities
    - Social connections
    - Events
- **Persistence** including knowledge infrastructure and services, which supports
    - Representation
    - Storage
    - Access
- **Application** including the architecture, services and development necessary for the implementation of the ER and CSG WeKnowIt prototypes.
- **Inputs and content** plays an important role in WeKnowIt:
    - Massive Web 2.0 referes to publicly available content in social media applications (e.g. flickr), which is used by the services in order to generate useful information through analysis and structuring (e.g. automatic construction of POI list for a region, presentation of images related to an ER event). This is the main source for generating Collective Intelligence within WeKnowIt. Please note that this content must not be necessarily generated by WeKnowIt and WeKnowIt users. WeKnowIt provides the techniques to analyse, present it and make it appropriate for usage in services and applications by the users.
    - WeKnowIt user contributed refers to content uploaded by the WeKnowIt system user and is used for content enrichment (e.g. user uploads a photo of an ER event) but also to invoke relevant services (e.g. automatic visual-based localisation). As mentioned the services might also depend on content of the first category to provide the result.
    - Non Web 2.0 training data. Some algorithms require manually annotated content to

work in a supervised learning manner. These data cannot be directly generated by social media sites (e.g. for speech processing).

The developed Collective Intelligence techniques in most cases exploit links, references and relations among different content items contributed by the users, thus differentiating from the legacy large scale data analysis techniques. Typical examples of such techniques are Flickr-based visual analysis, tag clusters extraction from massive user tagging and social media-based community detection methods. The integration of such different techniques originating from different intelligence layers could potentially leverage Collective Intelligence within diverse usage scenarios.

However, the Collective Intelligence approach proposed in WeKnowIt moves a step further; instead of a mere concatenation of the different layers intelligent methods, it imposes a combination of different intelligence layers within the architecture of some of the developed techniques. Multi-modal analysis is often exploited to enhance the results in each intelligence layer. For example, Mass Intelligence tag clustering results are improved by using Media Intelligence visual analysis features when building graph clusters. As a result the produced clusters are evaluated as more coherent, since they incorporate cross-domain knowledge.

Furthermore, the added value of Collective Intelligence is also evident in the integration level, where the different techniques are combined to produce better results in each case. For instance, evaluation results in Media Intelligence show that geo-tagging through visual and tag analysis yields better localisation results, while Flickr annotations can be improved through automatic localisation and tag recommendations emerging from massive user generated content or the user's social network.

By this Collective Intelligence approach, user generated content items can be analysed by Media Intelligence techniques, resulting in knowledge enhancements about them, as is the case for deriving routes from spatio-temporal analysis using the VIRaL tool. At a higher level, Mass Intelligence techniques are able of analysing massive Web 2.0 users feedback, aiming at extracting implicit knowledge about more generic situations, as is the case with ClustTour and its ability to identify interesting time periods or areas in cities. In overall, the WeKnowIt Collective Intelligence methodology is able to produce enhanced results by:

- exploiting large-scale user contributed content
- combining different layers in building Collective Intelligence techniques
- fusing results from different intelligent layers

In the above context, a series of 7 prototypes have been developed in terms of the two case studies, while another 13 prototypes and demos were independently produced. The following prototypes and demonstrators are available for use and can be accessed through the Results page of the WeKnowIt website[2]:

[2] http://www.weknowit.eu/tr

- Personal Intelligence
    - **Sparks**. Browse visually the hidden knowledge of RDFa documents through XHTML overlays.
    - **Attention-Streams**. Visualise and understand your interests while performing online activities.
- Media Intelligence
    - **Visual Retrieval and Localization Tool**. A content-based image search engine.
- Mass Intelligence
    - **City Exploration by Use of Hybrid Clustering (ClustTour).** Explore new places and events by use of an interactive map and photo clusters centered on Points of Interest.
- Social Intelligence
    - **dgFOAF -** Representing Distributed Groups in FOAF Profiles.
- Organisational Intelligence
    - **SemaPlorer** -Making Sense of Collective Data
    - ER Log Merging and Management (**WERL**)
    - **CURIO**: Collaborative User Resource Interaction Ontology
    - **Veracity**: Modeling and Proving Trustworthiness of Web Resources
    - **STEVIE -** Collaborative Creation of Semantic Points of Interest as Linked Data on the Mobile Phone
    - Multimedia Metadata Ontology (**M3O**)
    - **Mobile Facets**. A mobile application to access a large, distributed data set of different social media sources.
    - A Model of Events based on a Foundational Ontology (**Event Model F**)
- ER use case
    - **Emergency Response application**
- CSG use case
    - **Travel Planner.**
    - **Mobile Guidance** during trip.
    - **WeKnowIt Image Recognizer**. Mobile app that recognizes landmarks from taken photos.
    - **Fannr**. Geo-locate and tag your Flickr images.

A web page where available APIs from all partners are listed has been developed (http://mklab.iti.gr/wki-apps/). This page (Figure 2) describes in detail how the developed WeKnowIt services can be accessed and tested and provides links to responsible partner and person for further information and communication. The objective is to publicly provide to interested users the WeKnowIt services, but also to serve as a “one-stop” site where WeKnowIt specific technologies are presented and can be accessed.

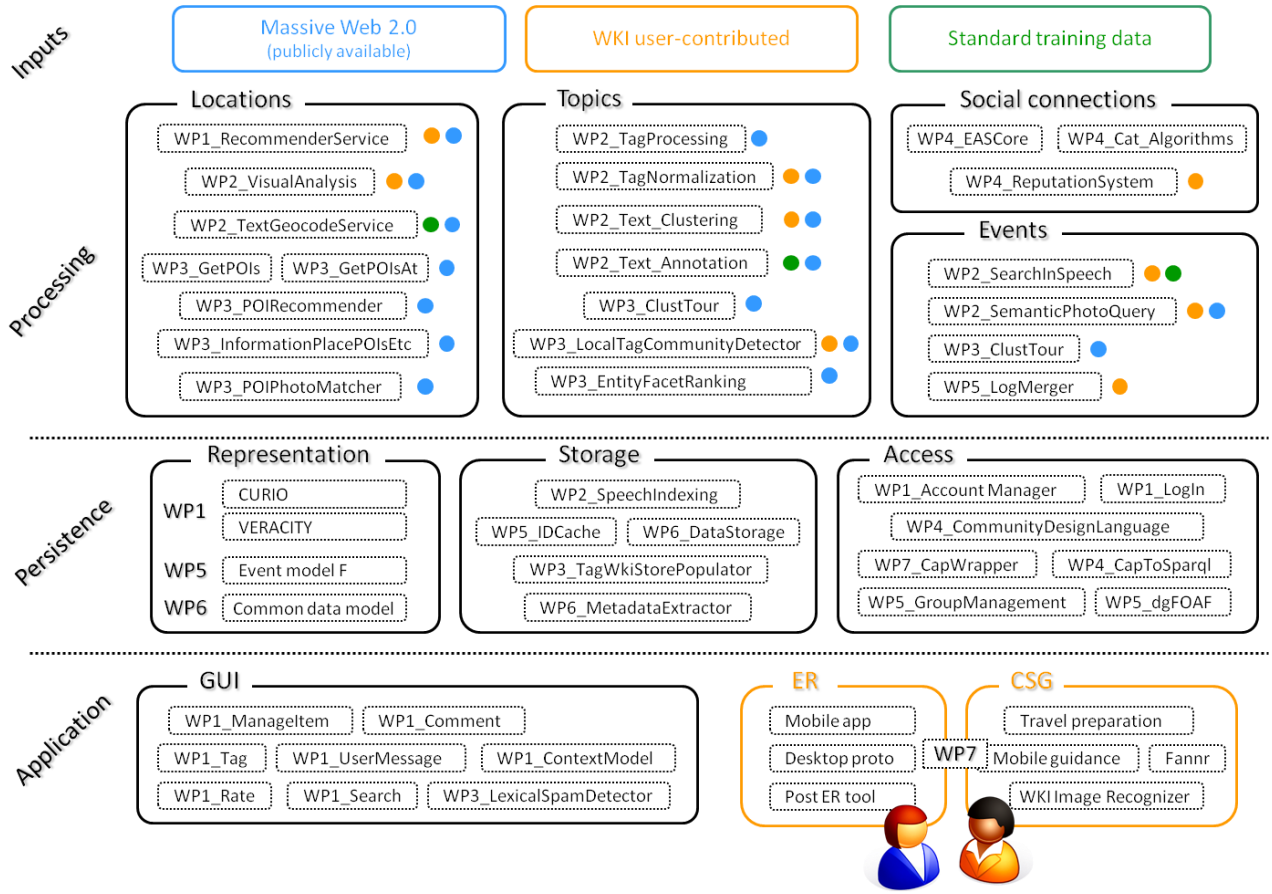

Figure 1 Overview of WeKnowIt services and various aspects.

This webpage currently includes the following publicly available APIs:

- **Attention Streams**: A semantic real time attention tracker
- **Speech Service**: WAV files processing, speech recognition and indexing, search for keywords in speech
- **Social Network Partitioner**: Fast partitioning of social networks into cohesive subgroups
- Text Classification: Allocate topics to text based on the text's similarity to a predefined model.
- **dgFOAF**: Manage user memberships for any given dgFOAF social group
- **VIRaL**: Issue queries with images and get list of similar images, the estimated location of the query image, frequent user tags and a list of possible landmarks depicted within the query image.
- **POI-photo**: Returns a set of photos depicting a POI, given its ID.
- **ClustTour**: Access the results of the hybrid image clustering method and the results of the cluster-based event detection method.

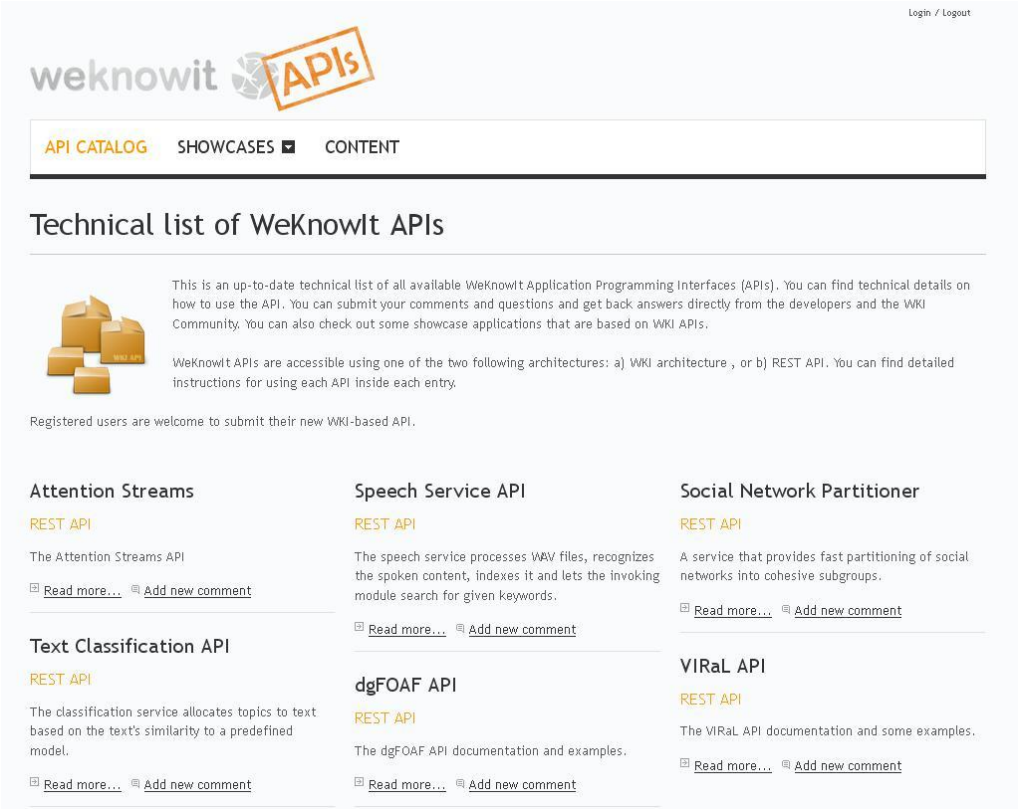

Figure 2 WeKnowIt APIs web page.

Due to the heterogeneous nature of the tasks addressed within WeKnowIt, the partners collected custom fit data that helped them solve the task at hand. The different used datasets do, however, come together through the integration of the corresponding services in the prototypes for the two different use cases. The task of collection, exchange and repackaging of multimedia content (text, images, video, and speech), user and user interaction data that was used within both use case scenarios as well as for the evaluation of research performed within the scope of the project, has been summarised in a single Content web page[3]. The content sources described therein enable semantic content analysis and metadata extraction, while user data enables social network analysis and trend detection facilitating the extraction and generation of Collective Intelligence. The interested reader in acquiring any of these datasets can come to contact with the responsible person indicated in each case.

In promoting user awareness the consortium was engaged in producing a video showcase, containing a presentation of the WeKnowIt project, including the final results, integration efforts, prototypes for the Emergency response use case and the Consumer Social Group case study, together with the independent tools developed within WeKnowIt. To this end a series of videos have been produced which can be publicly accessed and viewed (http://vimeo.com/weknowit). Also, a browse-able DVD has been produced, containing the different versions of the video, as well as chapters containing the independent WeKnowIt applications demonstrations.

---

3 http://mklab.iti.gr/wki-apps/content

During the project lifetime the consortium has produced 57 deliverables, 42 of which have been publically released and can be downloaded from http://www.weknowit.eu/deliverables. The deliverables describe all the new developed techniques within each one of the Intelligence layers, as well the architecture and integration of the WeKnowIt system components. Also the implementation of the two use cases and the evaluations that took place are also therein presented.

Finally, partners were engaged in 46 dissemination activities and have produced 79 publications, which are analytically described in section 4.2 of this report.

Personal Intelligence

Personal Intelligence dealt with the organization and presentation of information available in the form of videos, images or existing documents (e.g. Web pages, blogs, etc.), as well as semantic knowledge, which is derived from the processing of the former, adjusted to the preferences and current circumstances of individuals who have different informational needs and access requirements. The current state-of-the-art of Personal Intelligence upload and access was researched, understanding the tasks that Personal Intelligence should support and the interaction modalities involved, so to derive User Interaction Models to support the design and development of WeKnowIt applications. The state-of-the-art analysis focused mainly on personal intelligence methodologies and functionalities like, multimodal interaction interfaces, tagging and annotation and searching and browsing.

Moreover, work on interaction design has been conducted: taking into account the requirements derived from the WeKnowIt application scenarios and the related state-of-the-art, for each requirement one or more detailed interaction models were devised, with particular attention for the interface design. In particular research has been conducted on the state-of-the-art of personal intelligent methodologies and functionalities, such as Mobile Positioning, Push and Pull Information. Research has been carried out on high level requirements for Multimodal interaction, Personal information upload and access, Context modeling, Emergency response, Travel planning

The identification of possible tasks to be supported was undertaken by comparing the state–of-the-art with the gathered requirements and for each task the variables that may influence the interaction were outlined.

The Personal Intelligence layer also contributed to the building of Collective Intelligence by designing, developing and testing the methodologies and technologies for Personal Intelligence management. Since Collective Intelligence emerges from content available in different formats: videos, images, text messages or existing documents (e.g. Web pages, blogs, etc.), the content that is submitted by the user can either be already enriched with metadata (either implicitly or explicitly) or can be simple data. This content and relevant content submitted by other users together with

available metadata is then processed by Personal Intelligence technologies to extract hidden knowledge and enrich it with new semantic metadata contributing to the creation of Collective Intelligence. More specifically, research was conducted on the event-based representation. In cooperation with Organisational Intelligence and the Use Cases we have developed how to represent events and link them to the envisaged interaction models. We have also developed domain ontologies in cooperation with the Use Cases and have developed the Interaction Event model. As a result of the Personal Intelligence management, a graphical login mechanism has been developed (namely AsterID) 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. Personal Intelligence also developed the Ozone Browser within this framework. The Ozone browser is a client-side application that allows the user to visualise semantically aware objects in a web interface. It has a plugin architecture which means that it can be linked to other web services (e.g. Twitter) easily and also allows the user to summarise events, for example by viewing the location of the event on a map. This information can be also linked to the ontologies and web services present in the WeKnowIt system.

A further result was the creation of an Interaction Model for use within the Common Model framework designed for WeKnowIt. The Interaction Model describes the connection between the events occurring within WeKnowIt applications, the users of these applications and the resources that they provide to the system. The model was developed as an extension of standard ontologies.

The Personal Intelligence layer has implemented the following services to support information upload and access:

- AccountManager
- ItemManager
- LoginService
- TagManager

The following reference services have been also implemented:

- CommentService
- UserMessage
- RateService
- SearchService

In the following the Personal Intelligence strand of work has focused on three principal areas. Firstly, we have explored how to model users in a real time dynamic environment. We have investigated how to build a multi-scale model of user interests, which can represent both static long-term interests alongside short-term dynamic interests. This allows us to account for both levels when modeling users. We have also investigated how to model the context of the user. Again, context is multi-faceted property of the user and accounts for the users environment, technology and location. We

have built a flexible model, which allows WeKnowIt users to specify the context manually (e.g. if the user is in location X they are at the scene of an emergency) or to learn context through the sensor inputs (e.g. given the output of the accelerometer, is the user travelling or staying still). This allows the WeKnowIt system to account for the context of the user and deliver a more tailored user interface. Context is currently demonstrated within the system at the upload process. The users context is accounted for during the upload process and the user is presented with a smaller set of steps when uploading images if their network connection is limited or their battery is low.

Personal Intelligence has also examined how veracity (trust) should be handled within WeKnowIt. We have been exploring the notion of trust and how information is deemed to be trustworthy. We have also started looking at representations of how to model trust and to compute the trustworthiness of individuals within the system and external to the system. We have designed Veracity, an ontology for representing trust within the context of WeKnowIt. The ontology aims to represent the different facets of trust in contents rather than trust in entities. Particularly, it differentiates rational trust from social trust as a mean to represent the veracity of an information based on facts rather than social information. The model is more efficient than the existing models in the case of an unknown user adding information within the system. In this context, there is no information for asserting the trust in the asserted information based on social information. However, the trustworthiness can be derived from background knowledge that can be attached to the uploaded information thus providing an initial estimation of the veracity of the information. Our ontology is particularly relevant to the Emergency scenario since it is mandatory for the emergency personnel to identify relevant information in the context of trustworthiness.

Personal Intelligence also explored the notion of context within the boundaries of external user context. We have developed Attention Streams, a system which builds a picture of the users' attention through analysing their web browsing activities. The system then uses the information it has gathered to present the user with suggestions as to relevant material they wish to review (Figure 3).

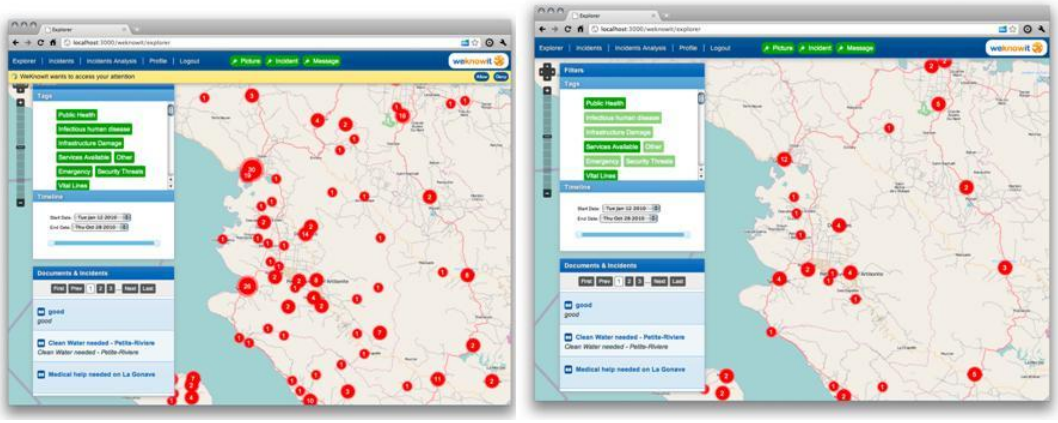

Figure 3 Attention Streams.

Finally we have also been exploring how to recommend locations, or Points of Interest, to the user on

the basis of their interests. Currently we are using Wikipedia to build models of each location and match these models with user interests in order to drive a flexible and relevant set of recommendations. After reviewing the relative literature a number of research directions were proposed as a result: adaptive itineraries, stability of user ratings, implicit vs. explicit interest inferences, short vs. long term interests, diversity and coverage. A survey of recommender systems approaches has also been made available. After examining the consistency of ratings and the process of re-rating in the context of user profiling and personalisation, the format of personalization was defined and implemented in the form of a POI recommendation service. The output is a ranked list of points of interest, locations, or events to visit based on a user's browsing behaviour. Relevance of a POI was based on cosine similarity between a user defined keyword vector, and the keywords for the Wikipedia entry for this item. Items were also slightly weighted with the ranking given by the Mass Intelligence faceted search technique in order to resolve ties between POI with similar keywords. The system for making the recommendations has been packaged up and is available as a web service of the WeKnowIt system.

Media Intelligence

Media Intelligence dealt with (semantic) analysis of multimedia content items, such as text, images, video or speech fragments and its efficient access by WeKnowIt users or in other words, Media Intelligence advanced current state-of-the-art multimedia content management and manipulation functionalities with respect to the potential WeKnowIt users.

Media analysis tools for all three modalities have been developed implementing novel research methodologies, particularly in the fields of visual content retrieval and localization, speech processing and Out of Vocabulary words detection, as well as textual events extraction and categorization.

Focus has been given on the impact of Collective Intelligence, i.e. the ability of a WeKnowIt users' group to exhibit greater intelligence on media analysis tasks than its individual members, in terms of improving analysis and retrieval of related multimedia content. Information from external sources (e.g. Wikipedia links of suggested tags) is also exploited in the process. Especially the visual retrieval and localization tool (VIRaL) developed emphasizes on the importance of the WeKnowIt collective intelligence, in order to achieve optimized retrieval results.

Significant efforts have also been spent on visual image retrieval and localization of digital still images through the production of VIRaL application (Figure 4), which facilitates efficient features and keypoints extraction and matching algorithms for classification to potentially identifiable objects, such as well-known landmarks. VIRaL is a photo search engine that can locate your photos on the map, it identifies landmarks and suggests tags to be added to a user photo.

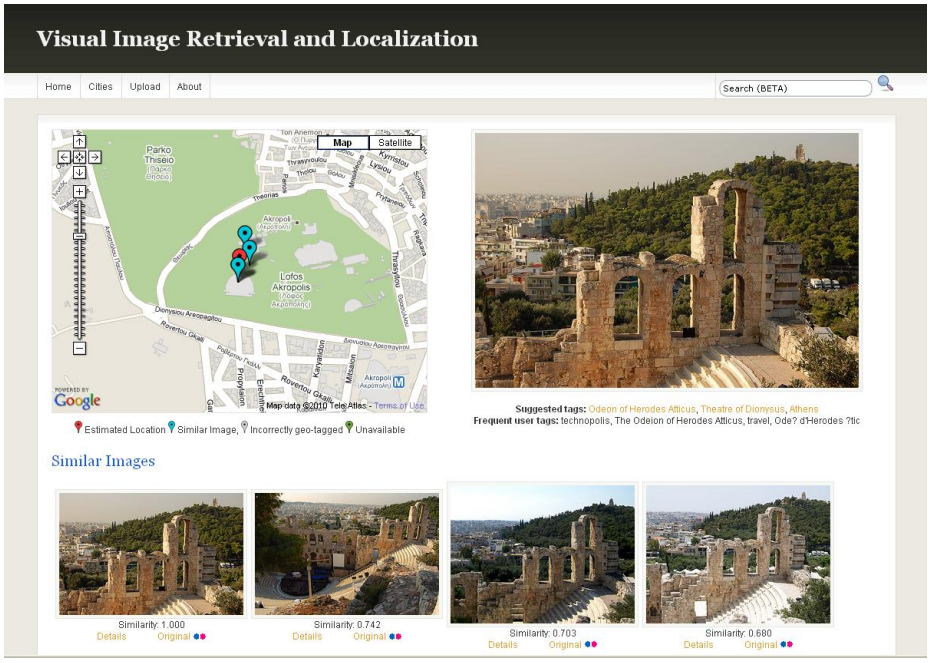

Figure 4 VIRaL application: Retrieval of similar images and localization of query image on Google Maps.

VIRaL finds visually similar photos within seconds, without having the user to provide any textual input about the photo. Its main innovation is that it locates any photo on the world map, it automatically identifies any depicted landmark, suggests tags and associates the content to Wikipedia articles. VIRaL's state-of-the-art visual similarity algorithms ensure highly accurate results within a very short time span. In implementing VIRaL a Visual Descriptor Extraction tool to be used in the pre-processing of training datasets had to be developed. In the same sense, research efforts had been spent on visual thesaurus production, as well as key-points extraction and matching algorithms. Research included clustering of still images to groups based on specific visual features and classification of these clusters to potentially identifiable objects, e.g. well-known landmarks. Then, a custom Flickr™ dataset collection and filtering was performed, in order to exploit data within development of the visual analysis tool. The collected Flickr™ image dataset within WeKnowIt contains now ~2.2 million geo-tagged images from 39 cities.

Another line of research concerning Visual Analysis was on photo clustering, based on associated tags of the photos, in order to get homogeneous groups of photos that relate to a particular topic. Some of the known drawbacks of tagging, such as ambiguity and lack of hierarchical relations were tackled, by integrating visual features into the clustering process and identifying visual centres.

Two visual analysis services (e.g. "WP2_VisualAnalysis" and "WP2_TagProcessing") have been implemented and integrated to the WeKnowIt development system. Tag processing, tag normalization and management techniques and methodologies were also developed and corresponding services were integrated to the WeKnowIt system.

The Media Intelligence layer of WeKnowIt offers a variety of innovative features and tools,

including also advanced textual and speech content processing. Textual achievements focus on extraction of important events from textual documents, by either automatically tagging texts or by indicting texts which can go on to further textual analysis (such as annotation). In the Text Analysis subtask, we have conducted research on geo-coding of social data (i.e. public forum websites), including examining the geo-temporal and geo-social nature of the data. A temporal aspect to events is considered in analysing the data and event “importance” is also measured. We also conducted advanced testing of Text Classification for Tagging of Emergency Types and developed Generic Annotations (based on Wikipedia/DBPedia Concepts). We examined also how text analysis tools are applied to data specifically relating to the Use Cases, and how issues arising, such as semantic under-specification in social data, can be addressed. Following this research, we have implemented a service to return proximate places (street names) to a location. Further work was carried out on combining user-driven semantic annotation and tag suggestions. Where the tags suggestions are influenced by the uploaded content and the user's associated semantic annotations (derived from the domain ontology). We have also implemented a geographic resource based on OpenStreetMap to ensure the techniques are generally applicable, rather than location specific, as well as implemented the contextual analysis and fusion work into the general media analysis techniques (and services). The above technologies and model have been also ported to Polish.

Speech analysis in WeKnowIt focuses on quick and efficient interpretation of raw audio and speech content by offering advanced techniques of Out-Of-Vocabulary words identification. Both tasks are proven to be very effective in tackling well-known limitations of traditional tagging and speech recognition systems. Specifically, we have performed research on advanced speech analysis techniques as well as research on the combinations of the standard recognizers with the search on the sub-word level. As a result we have developed OOV detection techniques. Finally, we have collected resources for training telephone-speech recognition and explored phoneme-lattice indexing techniques. More specifically, research efforts focused on the improvements of language modelling for LVCSR (large vocabulary continuous speech recognition) modules and on the combinations of state-of-the-art recognition techniques for telephone conversation with the search on the sub-word level. The technique applied takes benefit of the sub-word term candidate rescoring in word lattices. Advanced methods of the search in word lattices containing all possible alternatives of the pronounced words (as opposed to the one-best transcription) have been incorporated into the speech processing behind the WKI speech indexing and search services, enabling better accuracy in speech tagging. Further research and development of advanced speech processing methods were conducted, by combining vocabulary-based recognition with phonetic search. A new coding was explored which is based on temporal contexts in neural networks distinguishing the in-vocabulary words that are mis-recognized from out-of-vocabulary (OOV) words. Two relevant web services - WP2_SpeechIndexing and WP2_SearchInSpeech were built and integrated in the WeKnowIt system. In the following, these speech analysis services were improved by integrating novel techniques of feature extraction that are based on neural networks - bottle-neck (BN) features. These features

significantly improve the recognition accuracy. The speech recognition language models were subsequently updated to easily integrate lists of local names provided by the text analysis subtask. The new search in word lattices containing all probable alternatives of the recognized words provided a state-of-the-art performance and enabled both - efficient tagging of recordings based on a predefined set of keywords as well as the general search that actually matched the input query and the indexed speech.

The Media Intelligence layer also involved work on contextual media analysis and fusion. Regarding contextual media analysis the work focused initially on implementing a tag normalization process, i.e. matching the user assigned tags to domain ontology concepts. For the appropriate mapping between tags and formal descriptions, external sources of knowledge (such as WordNet, Wikipedia, etc) were exploited. Furthermore, string processing, matching and comparison functionalities were also employed in the process. In the following work was re-focused on implementing and utilizing tag processing information within the VIRaL tool. In this manner, its services were improved providing enhanced tagging information to the VIRaL tool and on integrating contextual knowledge in the actual process of the so-called "suggested tags" extraction. We also investigated clustering techniques with respect to the visual retrieval implementation of VIRaL tool, in order to utilize additional visual contextual information, available within the dataset of images, during the content retrieval phase. Finally, we implemented a service that gathers Flickr photos that are related to a specific ER instance. The photos are geotagged in a specific region, have been uploaded in specific time intervals and have tags that are semantically related to the ER-domain ontology concepts (the ER-domain ontology was developed within the WKI project). The interesting research aspect in the developed approach is the matching process between a tag and a domain concept. This is realized by mapping to each concept a weighted vector of terms (concept vector) that are derived from co-occurrence frequencies, use of WordNet and Wikipedia.

An equally important result in the Media Intelligence layer concerning contextual media analysis was the work on facet detection. The research builds upon an already established general framework for detecting facets for a given query. In the context of WeKnowIt, we have worked on the problem of facet aggregation from multiple trusted sources. In particular we worked on the fusion of facet concepts and relationships from Yahoo!'s GeoPlanet, Yahoo! Travel, and Wikipedia corpora. The developed methodology is integrated in the CSG case study, and also available as a stand-alone service through WeKnowIt service architecture. In addition, we have extended research in this area to the entertainment domain, which is of particular interest for image search. The research efforts have focussed on large scale, unsupervised de-duplication of concepts and facets from heterogeneous trusted (structured) sources, as well as concept variant detection. A typical example of the latter is a celebrity that goes by the name “Jennifer Lopez”, but also is generally referred to as “J-Lo”. We have also worked on expanding the coverage of Yahoo! faceted image search result exploration using fresh celebrity data. There was also work done on improving the Wikipedia based travel facets used

in the WeKnowIt travel services. This work has been extended for Yahoo! Image Search. Sports and event data was also added. Extraction of travel related data was implemented as a grid service for increased scalability.

In addition the fusion of information was also examined, discussing how to identify where facts concerning a given object (document) are related, i.e. (semantically) similar. The fusion processes which address the combination of reinforcing facts and resolution of conflicting inconsistencies are considered. The fusion of information is examined within the separate media analysis techniques and also the fusion of related facts derived from both textual and visual media. In order to improve results of these fusion techniques machine learning was used in order to determine a functional mapping between the input text and most likely toponym. Finally, all developed techniques are evaluated. In order to evaluate the work conducted within the Media Intelligence workpackage, we followed a thorough evaluation process, dealing with all aspects and media involved during the entire project's duration.

A final line of research in the Media Intelligence layer was social media intelligence. In this context, we have produced a Hybrid image clustering service. This service utilizes two different kinds of similarity, complementary to each other, in the sense that sometimes multimedia content such as a still image, may have missing user tags or it might be extremely hard to interpret it in a traditional visual way. Furthermore, different clustering algorithms and services have been integrated to the final version of the publicly available VIRaL tool.

All the above mentioned techniques and services from the three different modalities have been finally evaluated. All visual analysis methodologies and algorithms developed were evaluated and the results documented are very promising. To this end an integrated evaluation set of experiments has been conducted for both WeKnowIt scenarios using the VIRaL tool. Evaluation focused on both geo-location estimation and visual retrieval efficiency on collected WeKnowIt Flickr images datasets (both ER and CSG). The Text Analysis techniques were evaluated on the Ushahdidi Haitian data set which provided a real-world, directly relevant gold-standard. Finally, the SearchInSpeech and SpeechTagging methods were evaluated on the multimedia data set which is relevant for the scenarios dealt with in WeKnowIt. The additional value of the speech processing was quantified.

Mass Intelligence

The main aim of Mass Intelligence was to recognize and to understand facts and trends by exploitation of massive user contributions. In principle, it can be expected that the aggregation of data, metadata and user behaviour from and of a large mass of folksonomy members gives new insight that would not be possible by investigating the contributions at the individual level. To appropriately address the problems of Mass Question Answering, a number of challenging and fundamental problems was first considered, including self-organisation of data and knowledge in

large-scale social environments, principles of folksonomy growth and evolution and understanding of global user interaction patterns. Advancements in Mass Intelligence also included tackling of the problems of mass interaction feedback, mass classification and clustering, and mass evolution analysis. Specifically, techniques were developed for the 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, and on identifying communities of related tags in large folksonomies. Another approach to improve search experience in large-scale dataset is discovering facets and ranking them based on relevance to the user. Furthermore, hybrid clustering mechanisms have been developed, that take into account input from different modalities, such as visual and tag features. characterization of social media by combining text features with spatial knowledge, event and landmark detection in social media content, user and tag activity analysis over time in social media environments, faceted exploration of image search results, automatic localization of Wikipedia articles, and dynamic analysis of cluster evolution for burst detection.

In the Mass Question Answering line of research, after thorough analysis of the requirements stated by the WeKnowIt system, research focused on analysis of tags, questions and answers together with users and connections between them. Specifically, a first goal was the refinement of community detection techniques on tag networks. Then another line of research was the implementation of an automatic Spam Detection scheme suitable for Question-Answers systems. This work resulted in proposing and implementing the following two services, a) community detection in tag networks, and b) spam detection in QA.

For the community detection, a new method was developed that extracts tightly-linked groups of nodes from large networks in a progressive manner. That was possible through processing the local graph topology (local edge clustering coefficients, etc.). The combination of community seed set detection methods (e.g. the SCAN algorithm) with local-expansion methods was investigated and improved. Comparisons between the new method and the state-of-the-art have yielded promising results. Additionally within the scope of this research, we have implemented a tag network graph analysis and visualization using Java JUNG framework.

The automatic Spam Detection service is based on a scheme suitable for Question-Answers systems in a low-level. It is based on a series of consecutive filtering mechanisms, ranging from simple list-based dictionary matching to n-grams language models, was successfully implemented and tested. Obtained results prove the efficiency of the offered implementation.

In the same discipline of Mass Quesstion Answering, we have also developed solutions in the following two areas: a) use of latent topics and b) measuring answer quality. Research on topical analysis of QA dataset was aiming to discover latent topics to group both documents and users

around them. Measuring of the answer quality based on the provided question was further extended from the previously created method of the ontology-based text categorization. As a result of the performed research in these two areas, following services have been implemented: a) latent thematic group detection, b) discovery of expert users for given topic, and c) measuring answer quality. The framework for latent topic detection and discovery of expert is based on the LDA model (Latent Dirichlett Allocation). Using this generative model, latent topics are discovered based on the whole content of English version of LycosIQ dataset. In the latent topic detection service, newly incoming document is classified against previously discovered latent topics. In the expert finder service, system determines expert users based on quality of their contribution to topics discovered in the input document. This research in latent topics will be continued to include time dimension and to examine how communities and topics evolve over time. The second area focused on measuring answer quality. For this purpose we have extended previous research in the ontology-based text categorization to discover which topics are both relevant for question and answer. Definition of topics and their hierarchy was taken from the English version of Wikipedia. Sparsity of information included in the question part was a hard problem for applying pure statistical similarity methods. This method was aiming to overcome this problem with the use of general ontology.

Another line of research related to Mass Intelligence was Mass Interaction Feedback Analysis. In this context we have been extending the textual content-based ads recommendation (current research) with massive user feedback. Methods utilized here include using lexical graphs, co-occurrence graphs to perform vocabulary expansion, maximization of word relevance and focusing vocabulary-related graphs for ad expansion to a specific topic (e.g. soccer, music, …). A fast update algorithm for stream of tags that uses random walk model was also implemented, which resulted in creating a prototype. Finally we have investigated several approaches to analysis input streams of feedback data for incremental update algorithms for graph clustering to detect structures in mass user feedback. A reference implementation of an analysis module for fast update algorithm uses restricted random walk clustering model.

The Mass Intelligence layer also included research on Mass classification and clustering. This task involved creating concrete services and solutions for the WeKnowIt prototype that were based on performed research and requirements from the use cases scenarios. Developed solutions include faceted navigation that uses facets ranking mechanism (clustering of results; delivered first publicly available prototype), text categorization with use of ontology, tag recommendation, points of interest (POI) recommendation and new meta-clustering techniques.

Regarding the problem Point-Of-Interest (POI) clustering and recommendation, we had created a POI co-occurrence graph based on co-occurrence statistics of POIs in the same page/section/subsection of Wikipedia and then several strategies for finding related POIs were tried out. Second research area included work on the topic of hybrid clustering of Flickr images based on both visual and tag similarities among images (in cooperation with Media Intelligence). Another

research focused on tag community detection based on a two-step process: (a) local tag seed set selection and (b) seed expansion. This work was extended by use of the SCAN community detection clustering to operate without parameters. Finally, we have worked on the creation of a disk-based graph storage and accessing framework.

In the same context we have created services for facet detection using Wikipedia and geo-tagging location-related Wikipedia articles. Work on automatic geo tagging of Wikipedia pages was carried out in order to improve the WeKnowIt travel services used in the CSG prototype. Another task included extraction of knowledge from Wikipedia. It was used for detection of facets for a given topic of interest. This provided input to the user consumer group case study. In addition, research was carried out on placing images on a map, using collective intelligence derived from (geo)tagged flickr images.

As part of our research on faceted search and the "Web of Objects", we have worked on the ranking of facets, based on collective intelligence that can be derived from multiple sources, such as a large corpus of tagged flickr photos, and image search logs. Based on a probabilistic framework build on top of co-occurrence analysis, that tagging and search behavior of users is analyses and deployed to rank a set of candidate facets, given a topic of interest. Though the general framework for detecting facets is limited to a number of a-priori defined domains, the ranking of facets for a given topic is in principle unrestricted to a certain domain. The research outcome is integrated into the general framework, and made available to the consortium through the API services, and is also used at the server side of the CSG prototype. The work on the faceted exploration of image search results was subsequently improved with respect to the grid based ranking pipeline. The new grid-based facet system was evaluated with the Yahoo! Image Search production team and launched into production after passing the quality evaluation.

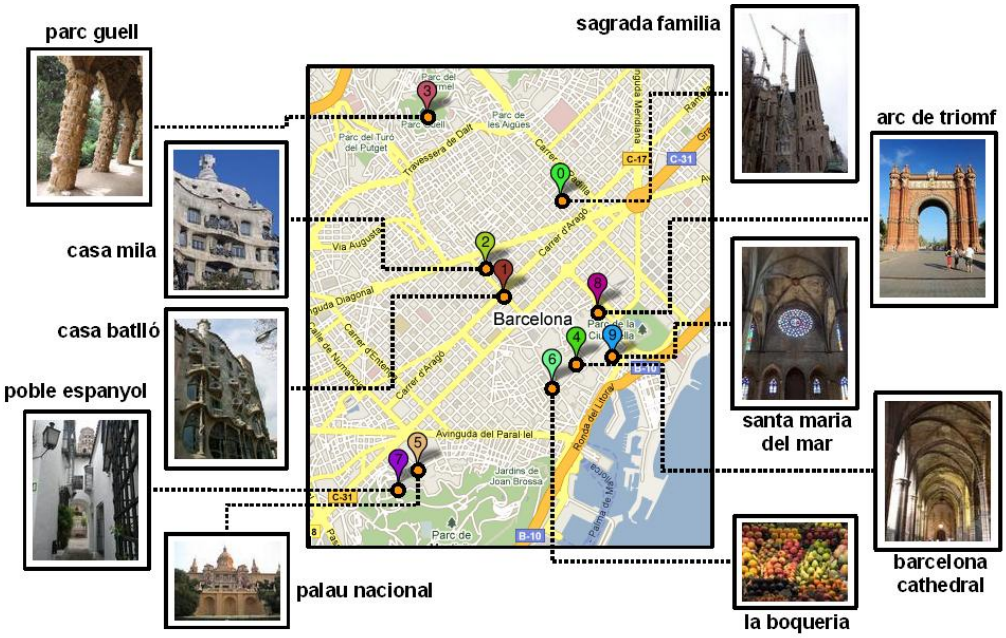

Figure 5 Top ten landmarks detected by ClustTour for Barcelona.

Under the scope of mass classification and clustering we have also worked on the automatic localization of Wikipedia articles, as well as on the identification of trending tagsets in Flickr data and on improving the Wikipedia-based travel services. Finallly, we have analyzed a large query log of 2.3 million anonymous registered users from a web-scale U.S. search engine in order to jointly analyze their on-line behavior in terms of who they might be (demographics), what they search for (query topics), and how they search (session analysis). We examine basic demographics from registration information provided by the users, augmented with U.S. census data, analyze basic session statistics, classify queries into types (navigational, informational, transactional) based on click entropy, classify queries into topic categories, and cluster users based on the queries they issued. We then examine the resulting clusters in terms of demographics and search behavior.

In the same scope of mass classification and clustering we have worked on named entities discovery using ontology, document categorization scheme based on domain ontology and meta-clustering approaches for flat and hierarchical clustering that leads towards topical navigation in a document set. Research on event detection in folksonomies and evolution of topics over time in social media was also carried out. Event detection has been evaluated for different created models on larger data sets to find optimal spatial and temporal parameters as well as granularity of detected events. It has been evaluated and a service was created for the WKI system. Evolution of latent topics and authors was tested on DBLP dataset that includes paper titles, authors and time information. Different distributions were evaluated to properly map topic activity over time.

A key research activity within Mass Intelligence layer concerned the application of community detection methods that were developed during the first two years of the project on the problem of photo clustering and cluster-based landmark and event recognition. The developed framework has been implemented in ClustTour, an online application demonstrating the results of the framework on a large photo collection of 25 cities (Figure 5). ClustTour has been extended in order to provide a spatio-temporal content organization layer on top of the photo clusters, and at the same time refine the detected clusters by taking into account the spatio-temporal distribution of photos in the dataset. In that way, ClustTour endows users with enhanced content exploration and browsing capabilities, and at the same time, it improves upon the quality of the presented clusters. The application is online at: www.clusttour.gr. We have also prepared a REST service for integration with WKI prototypes. Furthermore, work has been conducted on associating photos from Flickr with Points of Interest from Wikipedia, Wikimapia and OpenStreetMap. In this work, emphasis was placed on the retrieval of “long-tail” POIs, i.e. POIs that are not well covered in user contributed photo collections.

Mass evolution analysis was the final research task in the Mass Intelligence layer. It included extension of latent topics analysis with added time dimension, event detection in social networks and combination of latent semantics with geo location in social media. Work has focused on three areas: evolution of authors’ interest and latent topics over time, multimodal analysis of social networks for event detection, and combining tags and geo-location using LDA models for better content

organization. Modelling evolution of latent topics is an extension of previous work in latent topics analysis. It is extended with time dimension, to analyse evolution of (latent) topics over time and change of authors' interests using LDA-based (Latent Dirichlet Allocation) models. Work on detecting events combines analysis of tag distribution over time in social networks within geographically bounded location. It focuses on discovery "unusual" topics (hypes, outliers) in time that may indicate real-life events, like conferences, sports games, etc. Finally, research in GeoFolk exploits tag-location correlation in social media for better content classification, clustering and tag recommendation. In the same context we have been also investigating the dynamics of mass intelligence such as the change of tagging behavior over time. In the context of mass evolution analysis, we also introduced an approach for improved characterization of social media by combining text features (e.g. tags as a prominent example of short, unstructured text labels) with spatial knowledge (e.g. geotags, coordinates of social content). Further research was conducted on the problem of event and landmark detection in social media content. This culminated in the demo presentation of the ClustTour application in the ACM MultiMedia Technical Demonstration track. This demo was extended with the integration of spatio-temporal analysis of user contributed photo metadata in order to improve the photo collection exploration experience. Furthermore, a corresponding service has been developed and made available for the use cases. In addition, we have developed a framework for analyzing user and tag activity over time in a social media environment. The results of this analysis can be exploited also for event detection in social media, as well as for capturing trends and changing user interests over time. Finally, we have worked on the dynamic analysis of cluster evolution for burst detection. System is being developed to detect clusters related to predefined tag seeds. The focus was on the analysis of the evolution of the fuzzy cluster assignments.

Social Intelligence

The WeKnowIt Social Intelligence layer covers the processing of the social relations of online communities as well as providing a flexible authorization service for social applications. Social Intelligence used the relationships between members of WeKnowIt groups or communities as a source to enhance available information by the underlying relationships between people. It aligned itself to the WeKnowIt Emergency Response and Consumer Social Group case studies.

A main focus of this intelligence layer was the development of the community administration platform (CAP). Its objective is to provide a flexible tool to users and administrators to create and manage communities within the WeKnowIt system. In difference to existing approaches, the community administration platform can be plugged into almost any environment as it is realized as external service, allowing defining not only the usual policies and access rights but also the underlying model. The community administration platform is able to support in an emergency notification service within the ER use case. This service provides registered users to push and pull

information about personal status during an emergency. This should help the community to cope better and faster with information crisis during emergency. In the following we have developed an additional CAP2Sparql service for querying the CAP. An interface between CAP and the Organisational Intelligence distributed group management has been implemented that allows for integrating distributed defined groups with WeKnowIt access control. This service has become available in the final prototype and a RAM caching mechanism was developed to improve the performance of CAP access requests. Test runs showed that the performance for access checks might have improved by at least one order of magnitude.

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.

The Authorization Design Language is a formal language to build a community administration platform. An example realization of a structure buildable by the Authorization Design Language is depicted below.

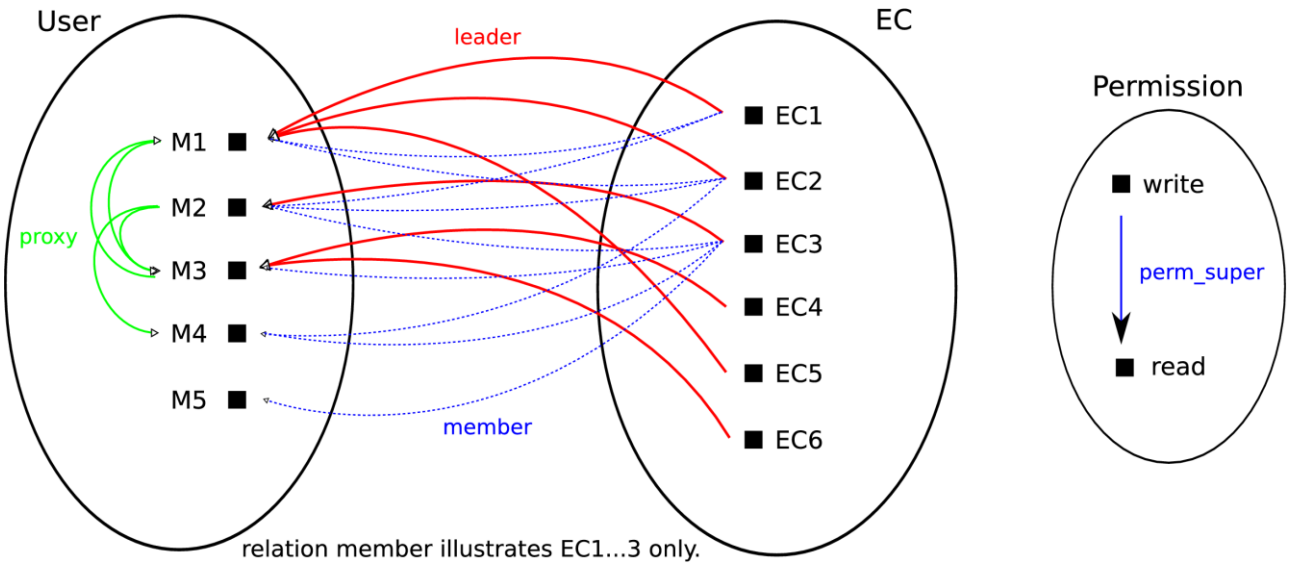

Figure 6 Authorization Design Language Use Case Illustration.

This figure shows an CAP example realization for the emergency use case scenario. Access rights can be modelled for the 5 members of the user base and 6 emergency cases in the system. A multi-relational modelling allows fine-grained control while enabling easy administration. Although not being able to describe all features of the Authorization Design Language for the sake of shortness,

one can see in the figure, that users (M1..5) can be assigned to be leaders of an emergency case (EC1..6) as well we ordinary members. The rights bound to these roles can be modelled very easy. Even proxies can be modelled by the green relation.

Apart from general characteristics of such an administration platform a new service has been defined: Emergency Notification Service (ENS). This service to the end user makes it possible that groups, subgroups or communities of predefined members can use this service to organize notification about the personal status during an emergency. The benefit of such a service is that no unnecessary phone calls or email services have to be used during such times. These services tend to be overloaded in times of stress. Thus the ENS provides a useful service not only to users by providing information within a group, but also helps to relieve standard communication lines.

A Community Member Life Cycle Model has been developed in order to describe user roles during community memberships. This structured life cycle model allows describing necessary roles, preliminarities and properties of roles and users. With these roles, a service can be defined allowing automatic role assignment/de-assignment to/from users, in order to allow users certain tasks. E.g. if an abuse of the system gets detected automatically, the user can be unassigned his roles. If a user, on the other hand, is very active and founds a community, he might automatically be assigned new roles, allowing him moderation of forum topics.

Efforts have also been spent on the design and implementation of community analysis tools. The community analysis tool (CAT) is a software library for the analysis of social networks. The objective of the tool is to provide social network analysis both for the other work packages and for end users in both use cases. The community analysis tool consists of two parts. The first is a framework for the implementation of analysis functionality consisting of the necessary data structures to represent social networks and the second part consists of algorithms to analyse these networks. The set of algorithms implemented in the ICAT are centrality measures, clustering algorithms and network statistics. A prototype of the initial community analysis tool has been also delivered. 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. In order to develop this tool we have investigated state of the art in social network analysis for social network sites. From leading and innovative social network sites the analytic capabilities provided had been collected. Furthermore, new analytic concepts were gathered from a literature review. The primary problem of this task was to handle very large evolving networks efficiently. Providing fast access and update mechanisms to graphs requires a lot of memory. To handle huge graphs a database-backed graph implementation has been developed. For complex analyses where instant calculation is not possible database-backed caches have been implemented. The final community analysis tool contains additional functionality for the spectral analysis of directed networks. Also we have implemented a module for the hermitean

eigensystem decomposition of directed, weighted networks. Further modifications of the Community Analysis Tool (CAT) lead to a new methodology in clustering methods. The main finding of the research is that the communication behaviour difference significantly between topics and that common expert identification algorithms are not suitable for all topics. Furthermore, a web-based demo application has been developed to show the capabilities of the community analysis tool.

Another task in the Social Intelligence layer was the visualization of community relevant information, aimed at providing a community browser that visualizes multi-layer networks and allows browsing them. Initially we had implemented a generic network browsing platform, based on which the WeKnowIt community browser was finally implemented.

A key line of research within Social Intelligence was the development of community services exploiting the cross-usage of intelligence. Analysis focused on scalable algorithms for community detection in networks. This capability of inferring social groups from huge communication networks is important for the Emergency Alert Service (EAS) – a prototype application for WeKnowIt's emergency response scenario. The EAS is a location-based service that activates nearby members of the social group of a victim in the case of an emergency. It is a mobile service that works as an emergency call agent, informing social contacts (friends, family, and colleagues) and public authorities about the emergency situation. A person in need activates a virtual button on his smart phone. The EAS checks the social network of the person in order to find persons which are spatially close enough to provide help and socially close enough to be expected to offer such help. We expect this approach to circumvent the so-called bystander effect. The service layout was modified twice to offer a more privacy-aware approach. Current geo-position data and routing information is provided. The service has been designed privacy-aware and dynamic: Friend lists are calculated automatically with behavioral methods. Friends are only alerted when physically close enough to provide help.

The Service of "Automated authorative role assignment" (AAR) was another cross-intelligence service. This service can be used to detect social roles in a community (e.g. the community leader). Social roles are associated with authoritative roles and automatically assigned or retrieved from the user. A user, who has been assigned such a role, can be rewarded with further responsibilities like assessment of validity of other users content.

Further work in cross-usage of intelligence has resulted to the development of two more Collective Intelligence Services, i.e. the WKI reputation system and a multi-modal recommendation engine. Both these approaches take into account results from other Intelligence Layers services in order to rate or recommend content items or users accordingly.

The Reputation Service describes a way to formalize the assessment of user generated content through explicit reputation. We consider reputation as a quantitative value that explicitly represents a property of an entity of interest, like users and content. It is used especially in the ER use case scenario for collecting and analysing the reliability of user-generated content. Since this content can

be the basis for important decisions made by ER personnel quality considerations apply especially in this case. The reputation system therefore provides its reliability scores for users and content items based on intelligence provided by the Personal, Mass, Social and Organisational Intelligence layers. Namely, it incorporates the Social Intelligence Community Analysis Tool, the Personal Intelligence Data Models and User Interfaces and the Mass Intelligence Spam Detection service.

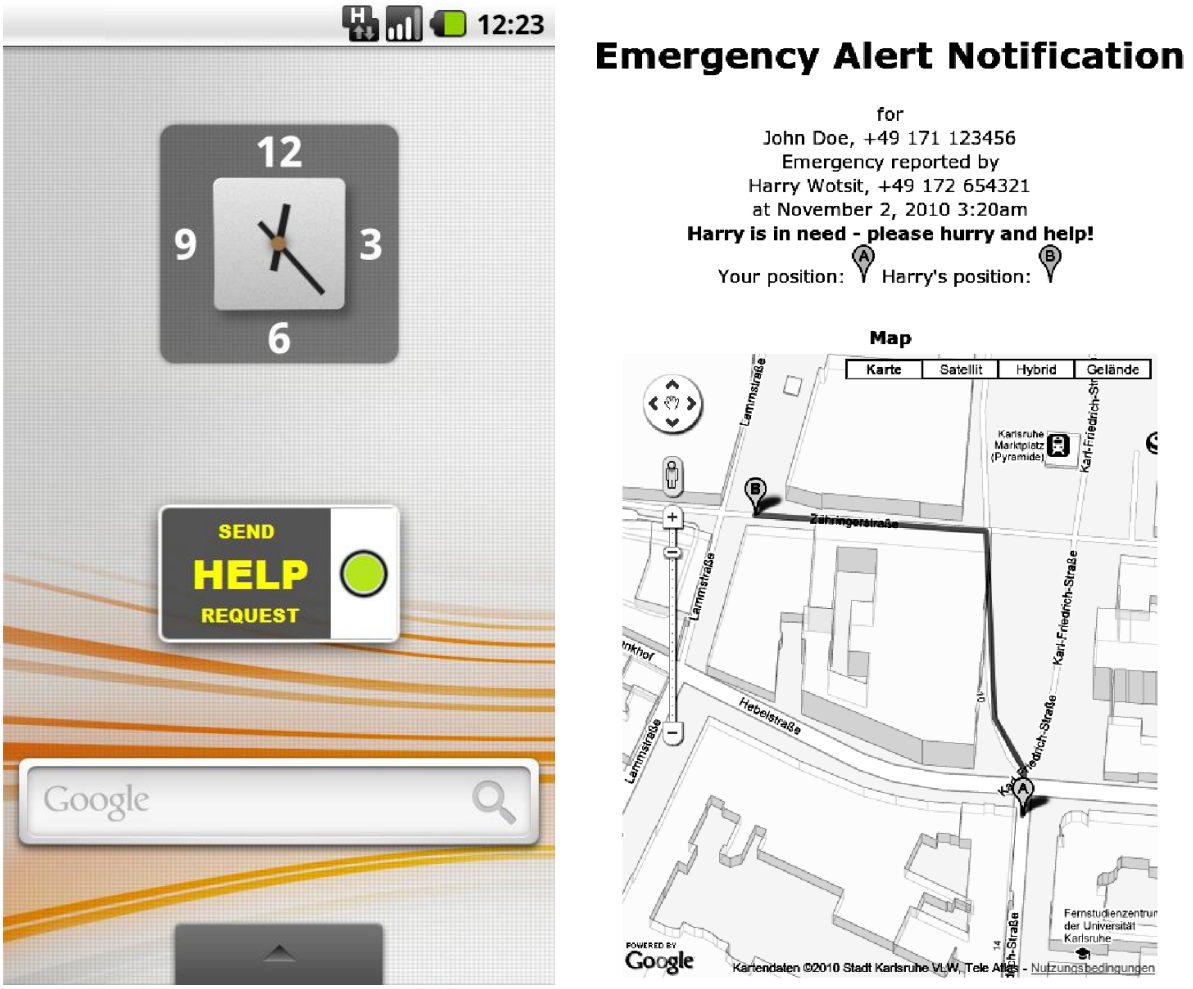

Figure 7 Emergency Alert Service: Homescreen alert widget on an Android smartphone (left) and emergency notification (right).

The User Recommendation service based on Cross-Usage of Intelligence exploits the natural multi-modality of social media and considers jointly intelligence from different modalities to output recommended groups of users. The users in each group are related with each other with at least one of the following ways: i) via their social networks, ii) by using related tags, iii) by uploading visually similar photos, iv) by any combination of the above. Such user groups can be exploited both in the ER and CSG scenario for identifying users that exhibit similar activity in a specific context. The recommendation service assigns relations between users based on intelligence results provided by Media, Mass and Social Intelligence.

Organisational Intelligence

In contrast to Personal Intelligence, the Organizational Intelligence dealt with the sharing of knowledge between the individual members of an organization. As a consequence, the role of Organizational Intelligence was to bring the right piece of knowledge at the right time to the right person of the organization in order to support decision making. This knowledge is not necessarily only produced by individuals, but rather by the interaction with Personal, Media, Mass, and Social Intelligence.

In the Knowledge Sharing line of work we have conducted analysis of traditional knowledge management systems and Web 2.0 knowledge management and investigated the information package (documents and media) provided by Sheffield City Council. In addition, we investigated in the modelling of relevant organizations in the WeKnowIt scenarios and usage of social network analysis for knowledge management. Then we initiated a discussion on a “common ontology set" to identify dependencies between the different layers of intelligence and to find data models to communicate information among the different layers of intelligence. This process resulted to a "Survey of Knowledge Management Systems" in the Emergency Response domain.

Organizational Intelligence tasks also analyzed knowledge management processes in professional organizations and Web 2.0 communities. As a result, a knowledge management methodology was developed. More specifically, an initial survey of traditional knowledge management and Web 2.0 knowledge management has been created. We first investigated in modelling policies for a networked definition of organisations and communities. The further, we started to develop a meta-model for such policies based on an analysis of different existing systems. The principal nature of professional organizations that have to make critical decisions such as an emergency response team have been also investigated. In the following we have developed a common model of events to allow for the exchange of event descriptions in the project. In a complex scenario where different, heterogeneous systems are involved such as the emergency response use case of WeKnowIt depicted in Figure 8 event descriptions need to be shared among the distributed systems of the emergency response entities. Examples of emergency response entities are the emergency hotline, police department, fire department, emergency control centre, and floating liaison officers. For example, the emergency hotline receives calls from citizens and sends descriptions of events reported by the citizens to the control centre. The emergency control centre is in charge of coordinating the emergency response entities. It receives event descriptions from the emergency hotline, processes them, and communicates event descriptions with the police department and fire department. In addition, the emergency control centre forwards event descriptions together with task descriptions to their floating liaison officers. Floating liaison officers are out in the field to report about a situation, verifying it, and documenting it, e.g., by taking photos and notes. In the case of a concrete incident this socio-technical system for emergency response becomes very active and many different pieces of structured knowledge such as events, tasks, and multimedia data with metadata are created,

combined, and communicated between the different entities involved. Humans think in terms of events and entities, thus events are a natural approach to represent human experience. The event model provided comprehensive support for all aspects of events such as time and space, objects and persons involved (constitutional aspect), as well as the structural aspect, namely mereological, causal, and correlation relationships.

The event model is based on the foundational ontology DOLCE. Existing solutions typically focus on the technical level of event management and processing of low-level signals and actions. With the event model F, we pursue an approach for describing events on domain-level and as happenings in the real world, respectively. These events are subject to discussion and interpretation by humans and by this very much differentiate from the low-level events. Domain-level events may be very complex and linked to a variety of modelling aspects and requirements. We have derived the requirements from existing models in various domains such as music, journalism, multimedia, news, cultural heritage, and knowledge representation. Identified requirements are representing participation of living and non-living objects in events, temporal duration of events, and spatial extension of objects. In addition, three kinds of event relationships are supported, namely the mereological relation (composition of events), causal relation, and correlative relation. The common model F also supports the experiential aspect, i.e., the annotation of events and objects with sensor data such as media data, and allow for different interpretations of events. Existing models almost fully support participation, time and space, and the experiential aspect. However, they substantially lack in the mereological, causal, and correlation relationships and event interpretations.

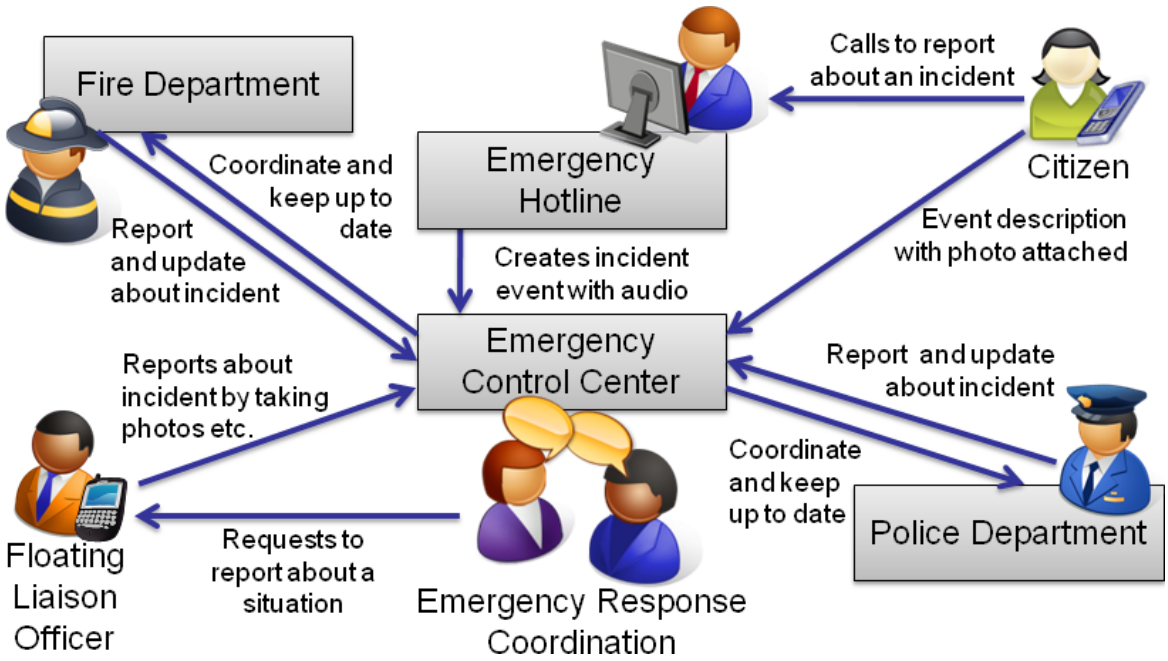

Figure 8 Sharing of event descriptions in the socio-technical system of the WeKnowIt emergency response use case.

For designing F, we base on the foundational ontology DOLCE+ Ultra Light and follow its pattern-oriented ontology design approach. More precisely, we use specializations of the descriptions and

situations (DnS) ontology pattern to provide formally precise representations of different contextualized views on events. This is required for modelling and describing domain-level events, i.e., the different aspects of happenings in the real-world. For designing F with respect to the requirements, specialized instantiations of the DnS ontology pattern have been defined for the participation of objects in events, mereological, causal, and correlative event relationships, event documentation, and event interpretation. Time and space are realized by the participation pattern and mereology pattern, respectively. Here, DOLCE-aligned ontology modules for temporal and spatial relations can be used. F distinguishes from existing models by its support for causal and correlation relationships and event interpretations.

With F, we can create and exchange sophisticated descriptions of real world events. For example, in the domain of emergency response one can model the participation of citizens in an emergency incident using an instance of the participation pattern shown in Figure 9. If the incident is a flood, one may speculate about the cause for the flood applying an instance of the causality pattern and possibly also using the correlation pattern. A flood may typically be composed of multiple events, which is modelled using instances of the composition pattern. Thus, for describing an event, different instances of the F patterns are combined, each providing a specific part of the event description. As there might be different opinions about the cause of the flood, there can be multiple instances of the causality pattern. To manage these multiple instances of the causality pattern (or other patterns), the event interpretation pattern is used to form different nexuses of the pattern instances and providing different views onto the same event. Thus, the interpretation pattern supports reusing parts of event descriptions on the level of pattern instances.

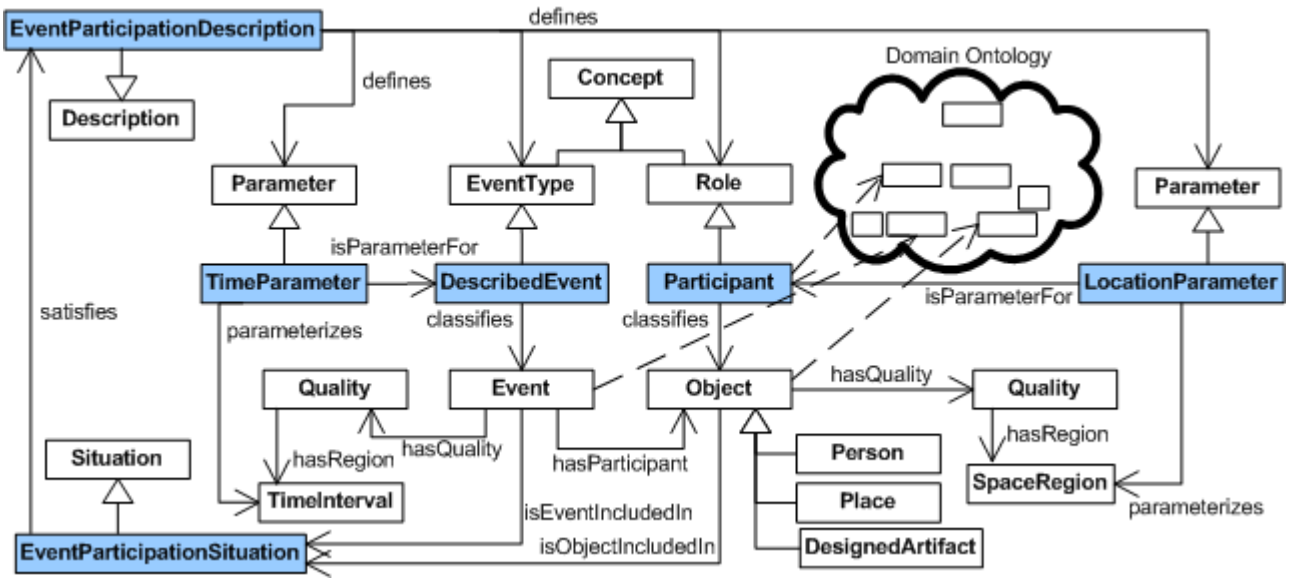

Figure 9: Participation pattern of the event model describing the participation of persons and objects in events.

In the WeKnowIt project, several event-based components provided from the other Intelligence

layers are dealing with event information. They need to exchange event descriptions like the one above. To integrate them and to effectively communicate event descriptions, they may leverage the formal model of events F.

The event model provided a flexible means for event composition, modelling of event causality and correlation, and allowed for representing different interpretations of the same event. By this, the event model can be used to represent the Collective Knowledge of the WeKnowIt users and the different interpretations the users may have about the occurrences in the world. The developed event model is based on an upper-level ontology and was developed in a pattern-oriented approach, modularized in different ontologies. It can be easily extended by domain specific ontologies. As such, it furthers serves as model for integrating the different WeKnowIt web services from Personal Intelligence, Media Intelligence, Mass Intelligence and Social Intelligence capturing the Collective Intelligence that is represented by the event model. Finally, as the common event model is designed independent of any concrete application domain it is capable to capture the Collective Knowledge of both WeKnowit case studies, i.e. Emergency Response and Consumer Social Group use cases.

This core ontology was named 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. It is employed in the SemaPlorer application for the faceted navigation in a very large, distributed and heterogeneous semantic data set. 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. We have also implementation an API for the Event-Model-F. Finally we analyzed W3C standardisation efforts for emergency response and compared it to the Event-Model-F. This W3C standardisation effort is the Emergency Information Interoperability Framework (http://www.w3.org/2005/Incubator/eiif/XGR-Framework-20090428/).

Continuing the investigation on the modeling of core ontologies and aligning core ontologies for the use in organizational intelligence have developed a Core Ontology for Multimedia Metadata (M3O) and also implemented an API for it. In particular, the four-step method for the integration of existing metadata formats and metadata standards into the Multimedia Metadata Ontology (M3O) has been further developed and documented. The M3O provides a generic modeling framework for multimedia metadata and provides among others support for the annotation, decomposition, and provenance of multimedia metadata. Subsequently, the alignment method has been extended for further metadata models such as ID3, Dublin Core, XMP, and Yahoo! SearchMonkey. We also developed a method for designing core ontologies such as the Multimedia Metadata Ontology (M3O) and the Event-Model-F. Finally we developed a method and tool support for automatically generating APIs out of ontology definitions.

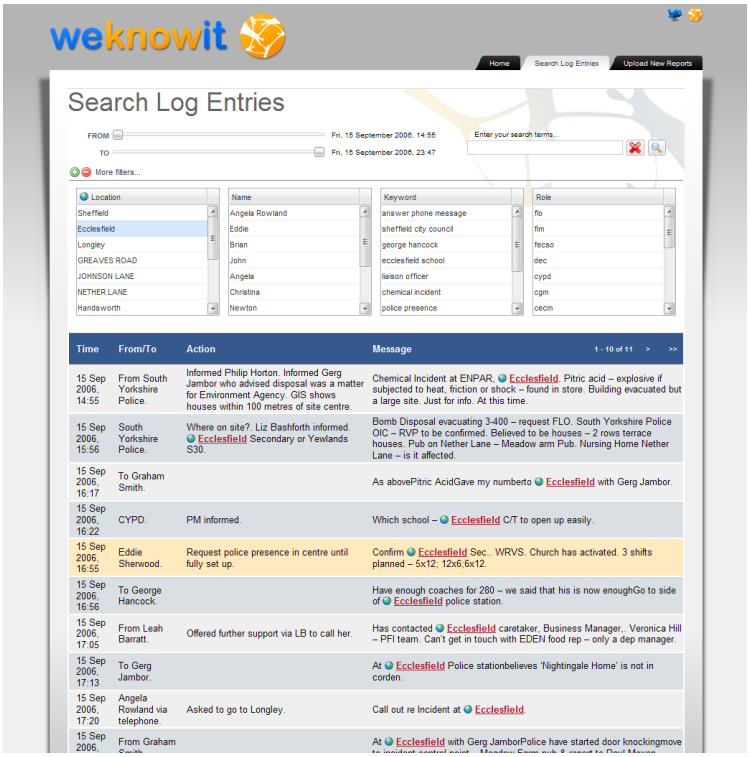

Figure 10: Snapshot of the WERL front-end main screen

In emergency response, the different professional entities such as emergency hotline, police department, fire department, and emergency control center use different systems with their own proprietary data models for events. 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.Using the formal Event-Model-F instead, these systems can commonly represent and effectively communicate event descriptions. The Organisational Intelligence Event-Model-F bases on the foundational ontology DOLCE+DnS UltraLight (DUL) and provides a set of ontology design patterns to represent the different relations of events such as participation of objects in events, mereology, causality, and correlation. It is used in the ER Log merging and management application WERL (Figure 10). This service supports the merging of log files created by different people and it provides search facilities over the merged log. It automatically merges different ER logfiles and represents them using the Event-Model-F. Furthermore, semantic information is extracted from their text and a slider-based time filter enables the examination of a particular time interval of the incident. Search options include free text, location-based, role-based and action-based search. WERL finally produces a unified view of an emergency event based on the partial log files produced by different members of the ER personnel. In addition to the service development, a prototype web front-end was developed for demonstration purposes. Online filtering capabilities based on lists of semantic entities were integrated. Namely, location and person names, ER-specific acronyms and important keywords extracted from the log text were presented to the user as handy filtering predicates. In addition, locations were automatically

highlighted in the text and linked to an online mapping service. Finally, it was possible to view filtered views of the logs based on provenance (who wrote the log) and time (when the log entry was recorded).

The Event-Model-F is also used in the SemaPlorer++ application, an extension of the SemaPlorer application, for creating and sharing event descriptions. The application allows its users to create event descriptions by clicking on a concept in the ontology on the left hand side, and drag and dropping it on the map. These events can be documented with textual annotations and pictures taken from Flickr.

In the same context we have also investigated the modeling of policies for a networked definition of organisations and communities (dgFOAF approach). We have created the joint dgFoaf authorization service. The Social and Organisational Intelligence authorization service is now able to use groups and links modelled externally by dgFoaf as authentication basis. The integration of dgFOAF group management into the WeKnowIt system and Social Intelligence rights management has been further pursued by aligning interfaces and functionality of the involved services. Moreover we have investigated means of quality improvement of user-generated content of the nonprofessional part of the organization.

Finally, we have developed and evaluated STEVIE, a mobile application for faceted search and exploration of social media (http://tinyurl.com/mobilefacets).

A third line of research in Organisational Intelligence was the development of a Knowledge sharing interaction methodology. We have developed a common ontology, considering the implications of Personal Intelligence and the use cases requirements. This process has led to the development of domain specific data models for ER (i.e. models for emergency type, participants etc.) and an Interaction Data Model (i.e. for modelling the interaction between the user and the WKI system). We have also gathered requirements for Knowledge Sharing and Reuse, investigating where knowledge sharing and reuse can be used to improve or satisfy the user requirements and development the WKI Knowledge Sharing Methodology. This work has involved much interaction with other work packages, in particular involving the UI (Personal Intelligence), the creation and augmentation of metadata (Media, Mass and Social Intelligence) and the pragmatic implementation of the interaction ontology (Integration). We have also developed the CURIO ontology (common data model) and its implementation in the prototype. The ontology has been generalised, divided into its core and extension parts, and the ontology specification has been published (http://purl.org/net/curio/ns#).

We also investigated how to preserve organisational intelligence about incidents by allowing users to post information about incidents from mobile and desktop devices. The background system uses semantic technologies to store the information submitted and allows for further sharing and reuse of the acquired knowledge. Preliminary analysis has been made on how to reuse the acquired intelligence in the post-incident analysis phase. We also developed a collaborative post incident

management tool and an experimental protocol for the evaluation of knowledge sharing. Primarily examining the degree to which the transfer and comprehension of information provided by the citizens, both individually and as a collective, is facilitated by the WKI system. The design of experiments in terms of Organisational Intelligence has focused in evaluating the benefit of using the CURIO ontology. Finally we examined the representation of trust in UGC with the publication of the Veracity ontology (http://purl.org/net/veracity/ns#).

The WeKnowIt System

The software framework created for the WeKnowIt System integrates loosely coupled modules produced by each intelligence layer. 40 services are currently deployed and integrated using Enterprise Integration Patterns. Among them, a data storage module is provided, that effectively stores and serves data of various kinds - multimedia files, relational data, Java objects and RDF triples. Within the WeKnowIt System data is exchanged using a Common Data Model, (based on the developed CURIO ontology). A custom Java-to-triples mapping mechanism is utilized.

Software solutions implemented within the WeKnowIt System are based on open-source software. Use case applications communicate with the WeKnowIt System via web services REST APIs, that exposes functionality of services integrated within the WeKnowIt System. Additional layers of abstraction of the WeKnowIt System, prepared per use-case, result in easy to use API tailored to the requirements of each application. Features of the publicly available REST API (i.e. its asynchronous nature) facilitate the creation of new end applications. This enabled the rapid creation of additional not initially foreseen applications, which exploit the WeKnowIt technologies, like the WeKnowIt Image Recognizer for mobile phones. The WeKnowIt Image Recognizer[4] is a mobile application that provides the user with detailed information on the location and name of a POI that she photographed. The application uses services developed in WeKnowIt in order to recognize Points-Of-Interest on the picture, determine its geolocation, and determine tags associated with this POI.

In order to build the WeKnowIt infrastructure we had first to gather the relative requirements. All WPs were asked for their expectations regarding the architecture. Also, the requirements resulting from services defined by WPs were taken into account. Special interest was put into discovering of the requirements of declared services in terms of hardware and performance. Based on the collected requirements, we performed various investigations and tests with a large number of mainly open source solutions suitable to fulfil the stated requirements. The resulting architecture is based on Service Oriented Architecture (SOA) approach, managed by an Enterprise Service Bus (ESB). Specifically, Fuse ESB (enterprise-ready ESB solution based on Apache ServiceMix) version 4 was selected as the final solution for service infrastructure. A group of accompanying standards,

[4] http://www.weknowit.eu/wkiimagerecognizer/

technologies and frameworks was also selected: OSGi (modular component model for Java), Apache Camel (message-routing and integration framework), Apache CXF (web services integration framework), JMS (messaging system), Spring (Dependency Injection framework) and other. All these technologies were tested for their compatibility.

Several architecture layers where identified:

- Business Process Layer, groups services from Services Layer into processes and exposes some of them via Web service endpoints.
- Services Layer offers functionalities used by higher layer. It uses DAO layer to access WeKnowIt Data Storage. Some services are exposed via Web service endpoints.
- DAO Layer provides access to the storages hiding the complexity and multitude of used storages behind an API.
- Storage Layer is responsible for storing resources (files, semantic and relational data).

This list was further extended with layers of applications that make use of the functionalities provided by the WeKnowIt System.

The hardware architecture has been selected based on the estimation of performance requirements of the use-case applications – both Emergency Response and Consumer Group Study.

In parallel we have defined the infrastructure required for development of the WeKnowIt project. This infrastructure consists of:

- Hudson[5], Continuous Integration server,
- Sonar[6], code-checking tool,
- Nexus[7], Maven repository manager,
- Mantis[8], bug tracker,
- WebDavs[9], document repository,
- SVN[10], source code repository,
- Integration environment.

---

5 https://hudson.dev.java.net/

6 http://www.sonarsource.org/

7 http://nexus.sonatype.org/

8 http://www.mantisbt.org/

9 http://en.wikipedia.org/wiki/WebDAV

10 http://subversion.apache.org/

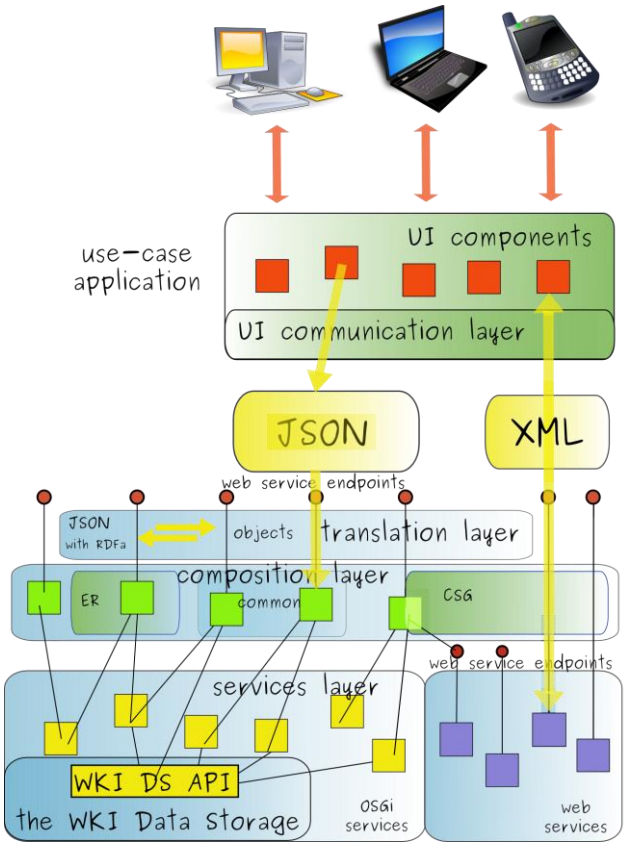

Figure 11 The WeKnowIt System architecture - with use-case applications.

We have also been maintaining a set of guidelines to help with WeKnowIt services development and use of development infrastructure. Guidelines were consistently updated so as to reflect the current status of WKI System and changes in the Common Model. In total 10 guidelines are available:

- Development Conventions (ver 0.1),
- How to install Java & Maven (ver 0.6),
- How to install the WKI System (ver 0.15),
- How to create Maven project (ver 0.9),
- How to upload missing JAR to Maven repository (ver 0.2),
- How to create OSGi service using Spring Dynamic Modules (ver 0.9),
- How to deploy OSGi bundles to the WKI System (ver 0.3),
- How to use Hudson CI (ver 0.5),
- SVN project layout (ver 0.2),
- How to run Berlin SPARQL Benchmark (BSBM) to test the WKI Data Storage (ver 0.2),
- How to map DTO to Curio graph (ver 0.1).

In the following we have defined a common data model. A list of objects was identified (it includes for example tag, document, user, tag, etc.). A set of ontologies describing these entities was established, and Java implementation of the ontologies was provided. The common model was used in all system prototypes. An integration environment was developed and updated regularly with new versions of services. Numerous issues related to technical side of integration (mainly OSGi, SVN

and Maven issues) had to be solved. Ad-hoc code reviews were performed during integration of services and found issues were discussed with responsible developers (i.e. problems with logging and exception handling). Code review activities were performed in order to discover the problems as soon as possible and avoid the spreading of bad code into next services developed. Automatic code checks (with use of Findbugs & PMD) were performed on Continuous Integration server on daily basis. All developers had access to code quality reports.

Another achievement during the Integration of WeKnowIt was the implementation of the WeKnowIt Data Storage (DS). A design of hybrid solution (triple-store and object database) was finalized and implemented. The implementation was preceded by extensive research and new requirements of the WPs were analyzed and discussed.

The final DS solution consists of File Storage (for storing multimedia files) and Knowledge Base (for storing metadata).

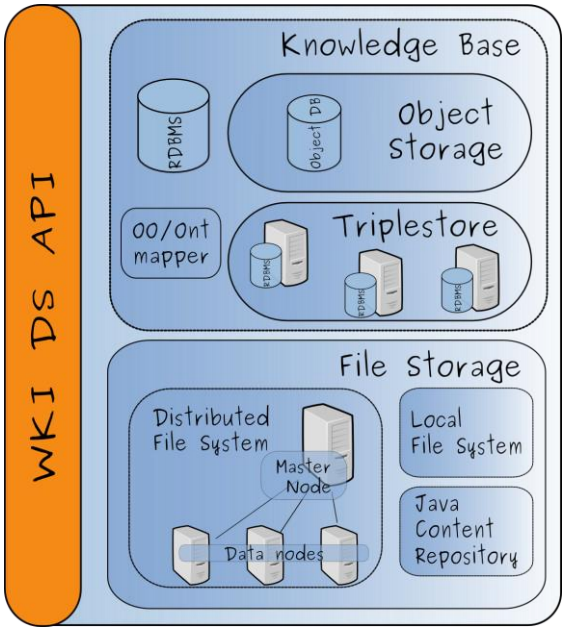

Figure 12: Elements of the WeKnowIt DS.

File Storage has a pluggable architecture and can be realized with different technologies. Three solutions are provided for storing files: filesystem, Distributed File System (DFS) or Java Content Repository (JCR). Implementations for the DFS and JCR were selected (Hadoop[11] DFS and Apache JackRabbit[12] respectively) and integrated with the WeKnowIt DS.

The Knowledge Base is a hybrid solution which uses triple store and object database (a RDBMS is also used to store some WKI DS specific data). This approach allows to achieve a unique combination of flexibility offered by triple store and speed offered by object database - a valuable

---

11 http://hadoop.apache.org/

12 http://jackrabbit.apache.org/

features for applications that demand both fast objects access (e.g. retrieval of user/image/event data) and complex search capabilities (i.e. SparQL entry point).

Solutions for each component of the Knowledge Base were chosen and integrated:

- triple-store - Jena[13]
- object database - NeoDatis ODB[14]
- RDBMS - MySQL[15].

Substantial effort was also put into development of the additional layer that facilitates the use of common model with the WeKnowIt DS. The API of the WeKnowIt DS evolved because of the new requirements of created services and common model initiative. The goal of the API has not changed. Its role is to shield the developers of services from the internal complexity of the WeKnowIt DS, provide all required entry points (e.g. SparQL entry point) and make use of the WeKnowIt DS simple and intuitive. The WeKnowIt DS was used in the prototypes of the WeKnowIt System and is used as a part of integration environment proving the compatibility of the WeKnowIt DS with the whole system architecture. Some effort was spent on benchmarking of the implemented solution using the BSBM benchmark suite. Some performance fixes were introduced based on the results and experiences gathered during benchmarking.

Finally, a web page where available APIs from all partners are listed has been developed. The WKI-APIs page[16] describes in detail how the developed WeKnowIt services can be accessed and tested and provides links to responsible partner and person for further information and communication. The objective is two-fold: on one hand to publicly provide to interested users the WeKnowIt services, but also to serve as a "one-stop" site where WeKnowIt specific technologies are presented and can be accessed. If you are interested in exploiting WeKnowIt technologies in collaboration with the project partners, instead of visiting several sites or downloading documents and software in different formats, you can find there a lot of the WeKnowIt services in a uniform format. For more detailed information regarding each service, the responsible partner and person can be contacted.

The WeKnowIt Use Cases

WeKnowIt established two use cases to demonstrate the wide applicability of its technologies and research activities. The Consumer Social Group case study provides the ability to extract meta-information from various content sources and user generated content within the system, as well as

---

13 http://jena.sourceforge.net/

14 http://www.neodatis.org/

15 http://www.mysql.com/

16 http://mklab.iti.gr/wki-apps/

enhanced publishing tools to support travel activities for single persons and groups. Based on the travel behaviour off all users and an analysis of their profiles, WeKnowIt information is extracted about the most preferred travel destinations and the trends and changes in user behaviour. This information is then used to make recommendations to the users and support the decision making process before (e.g. destination) and during the trip (e.g. restaurant).

On the other hand, the Emergency Response case study aimed to design, implement and deliver technologies and methodologies that enable citizens distributed across the region to participate in the monitoring of an incident or event. This methodology will benefit Emergency Response planners that will have real-time information available on which they can base their decisions and strategies, enabling them to better react to an Emergency. Thus, an emergency planner is able to find exactly the needed knowledge amongst all the available information and to selectively make this knowledge available to the citizens (e.g. information about open roads, information about relatives involved, etc.) in a largely automated way. The technologies in use also encourage and enable meaningful dialogue between the Emergency Responders and individuals, groups and communities.

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. 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. 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.

The work on the ER demonstrator[17] has mainly focused on creating the prototype by taking into account as many research functionalities as possible and additional requirements and feedback from the users. To respect the multimodal requirement and the need for context-driven interaction two

[17] http://nebula.dcs.shef.ac.uk/weknowit/

versions of the prototype have been created; a desktop interface and a mobile interface. The mobile interface has been implemented as a customised application for Android Smartphones. The desktop interface (Figure 13) has been enriched with new functionalities, such as a messaging service to instantly reach users on their smartphones, the possibility to submit incident reports both from desktop interface and mobile application (with automatic geolocation) and the capability of performing post-incident management searches and analysis. Newly developed intelligent services were integrated such as a tag analyser and a reverse geo-coding web service, services for managing users and groups, as well as services for the background architecture. A new functionality for uploading audio files was added to the mobile application, allowing users to upload messages that can then be listened to from the desktop interface (this functionality is ideal for emergency situations, to allows emergency responders to upload descriptions of the event or comment and to keep the recording of this intelligence for future analysis).

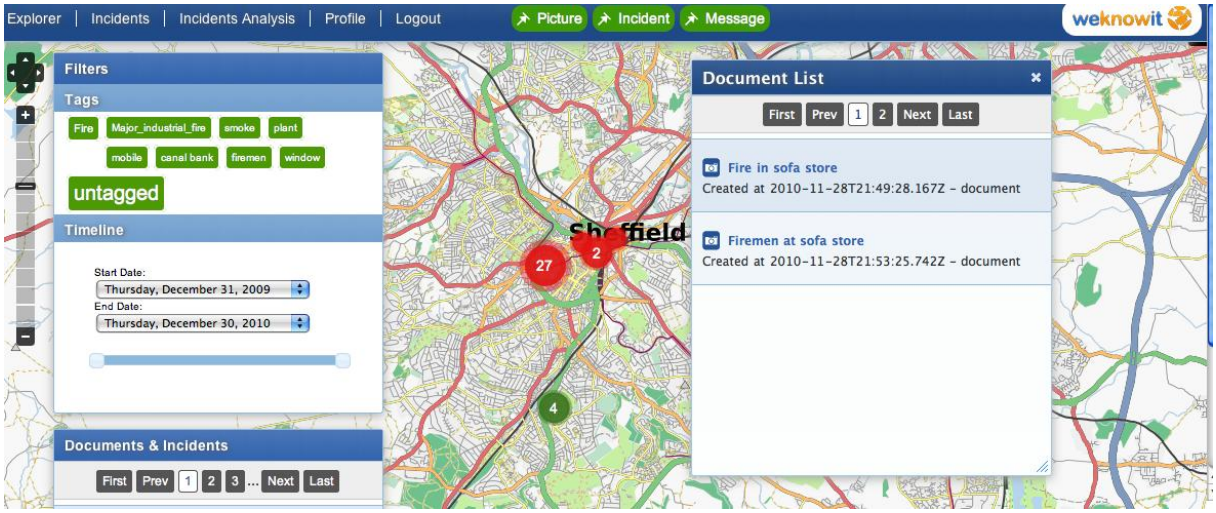

Figure 13: The ER desktop interface.

The evaluation of the emergency response application was undertaken in different sessions, following different methodologies and with different participants for each session. Comparing the results of the three evaluation rounds, the final scores are positive especially considering the prototype status of the application. No major usability problems were identified (Figure 15) but feedback was given on some aspects off the system that could be improved, and some improvements have already been incorporated in the system. It is very important to notice how the users appreciated the quality of the intelligent suggestions provided by the system, especially for the geolocation and reverse geo-coding system and the tag suggestions. During a longitudinal study evaluation was clear how at the beginning most users did not fully understand the reason for tagging or using automatic suggestions, but by the end of the study it had become evident that tagging the information using the automatic tag suggestions was a very quick means to easily retrieve the information at a later stage and communicate valuable information to the persons accessing the system (Figure 16).

Figure 14: ER evaluation exercise.

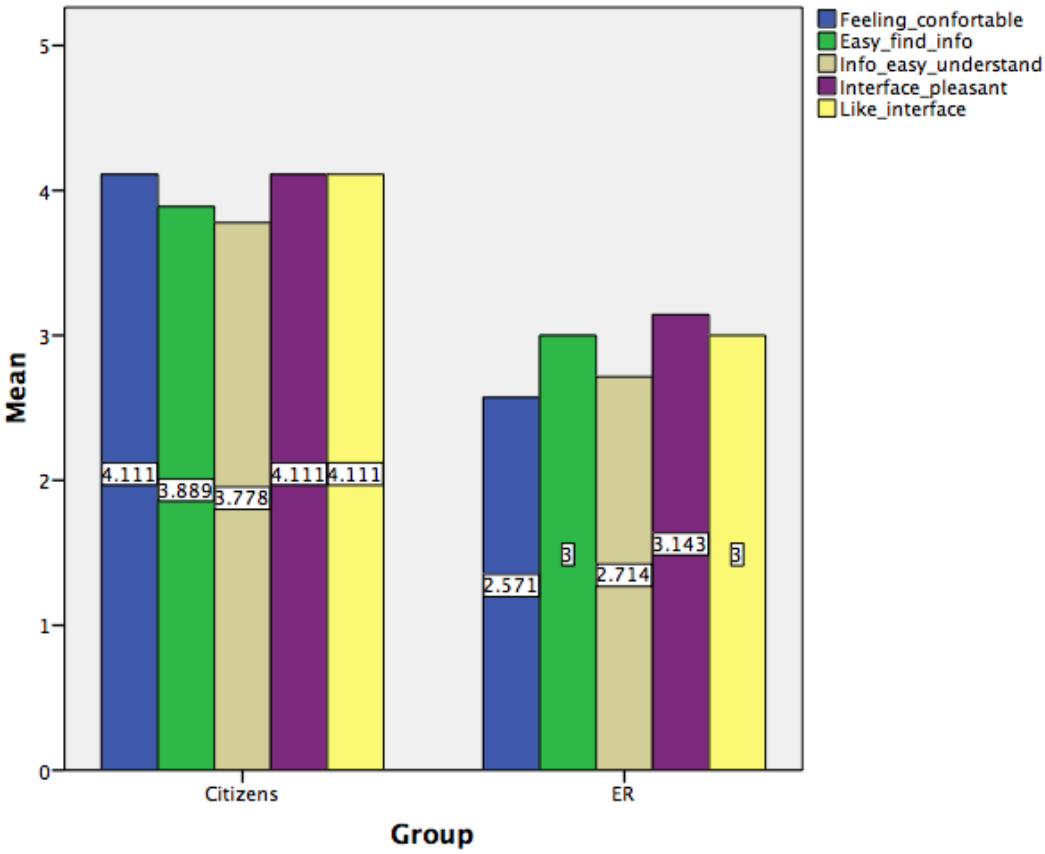

Figure 15: ER overall system usability.

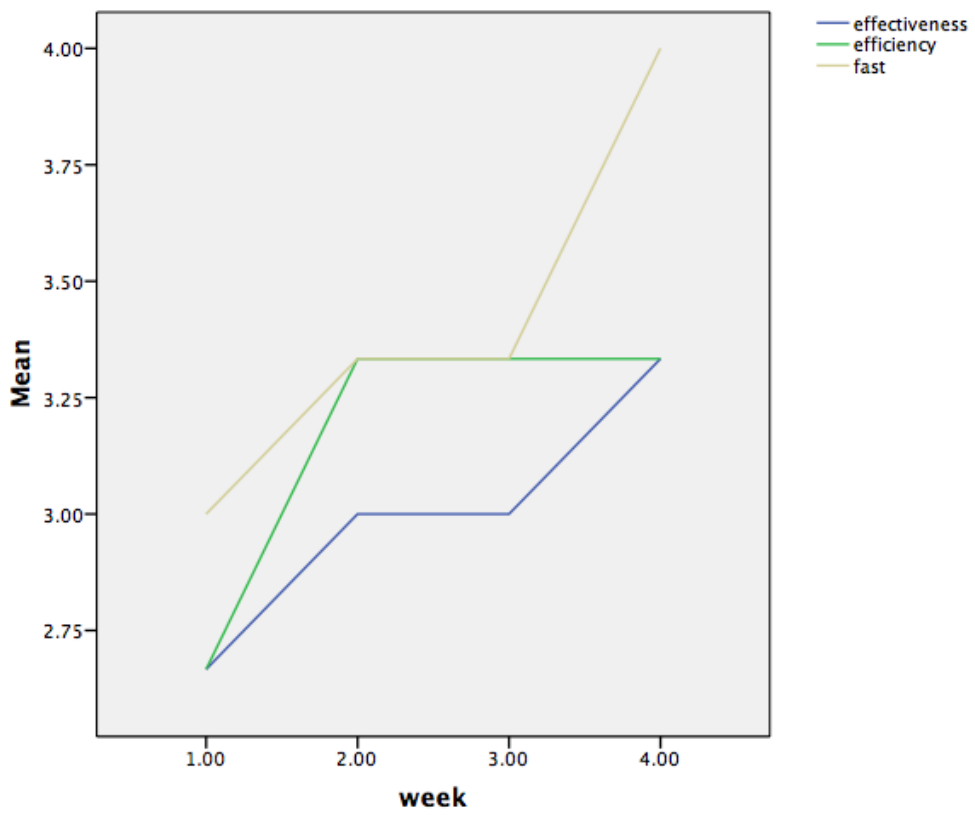

Figure 16: ER System efficacy and efficiency over time.

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 is about preparing the trip with a personal computer at home, and the second part is about supporting the group by means of guidance application running on a mobile phone while having the trip.

The prototype features a rich internet application[18] to display information related to the travel, search results based on Mass and Media Intelligence components, localization of points over maps to enrich the user's context, display of images related to the selected query, location, events 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. Various WeKnowIt services were also integrated, such as, point-of-interest recommendation and hybrid image clustering. Storing of bookmarks was implemented using the del.icio.us API, allowing for easier exchange of data between prototypes.

The second part of the scenario involves getting mobile guidance[19] during the trip. In this phase, the user gets useful and helpful information on-line on the mobile phone; besides he can take pictures

[18] http://weknowit.research.yahoo.com/csg/

[19] http://weknowit.tid.es/jsp/login.jsp

recording pieces of his experience. The prototype allows for a mobile version of searching for places, events and points of interest, providing detailed information of the places, events and points of interest found, making use of social and personal intelligence, as well as mass content analysis techniques in Wikipedia articles and other public sources. The mobile guidance also features enhanced map controls, display of events around the user, access to bookmarks taken during the pre-travel phase, and photo cluster visualization.

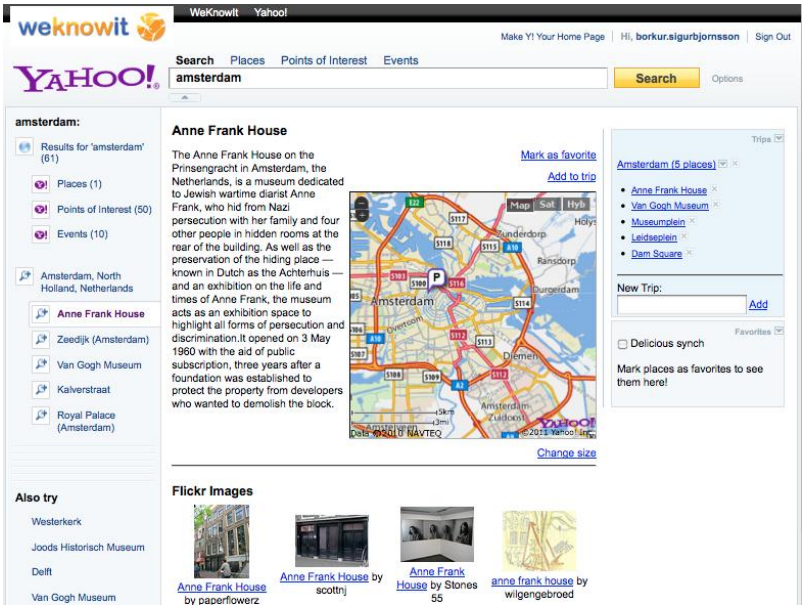

Figure 17: Travel planner interface snapshot.

In addition, a post-travel prototype, namely Fannr[20], aids users when annotating their photos after the travel. Fannr (Flickr Annotator) is a prototype application for assisting users in annotating their Flickr photos. It integrates several services from WeKnowIt project partners. Two types of annotation support is provided when annotating a single photo. 1) Tag recommendation, where the user is presented with a list of tags that they could consider adding to their photo. 2) Location recommendation, where the user is presented with several potential locations as to help them remember where the photo was taken. Both the tag recommendation and the location recommendation is based on data from different social levels: the user's own photos (personal intelligence); the photos of the user's contacts (social intelligence); everybody's photos (mass intelligence); and visually similar photos (media intelligence). Furthermore, for photos that have been localized the tag-recommendation is also based on geographically close-by photos. When the user has identified tags that they want to add to their photos or identified the location where the photo was taken, the Fannr prototype updates the annotation on the Flickr website using the Flickr API.

The CSG prototypes have been evaluated in two cycles: Their first version was evaluated during

[20] http://weknowit.research.yahoo.com/fannr/

December 2009 and January 2010 in Madrid (Spain) and Barcelona (Spain). The final version of the CSG prototypes was evaluated by independent users in three different places: Athens, Krakow and Barcelona. The CSG evaluation process regarding the prototypes indicated that there were no major usability problems identified (Figures 20, 21) but the evaluation has pointed out some aspects that can be improved in further development of the prototypes. One of the main conclusions was that the evaluation results depend on the location where the evaluation takes place. In hindsight this is obvious since the prototype is dependent on collective knowledge from social media sites and the support of big cities is much better than for smaller cities. The experience in three cities has been different but helpful to providing a more complete view on how the prototypes can be further improved.

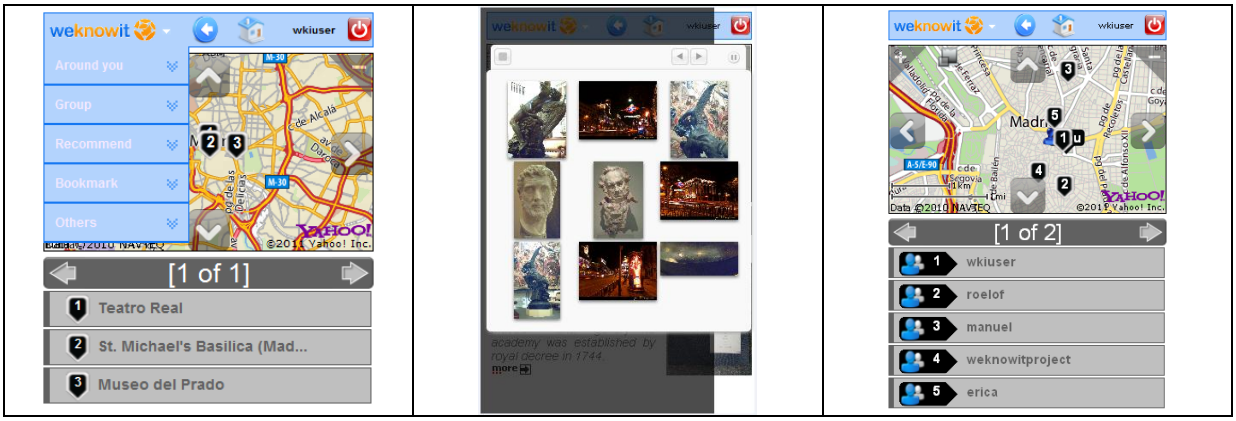

Figure 18 CSG Mobile guidance interface snapshots.

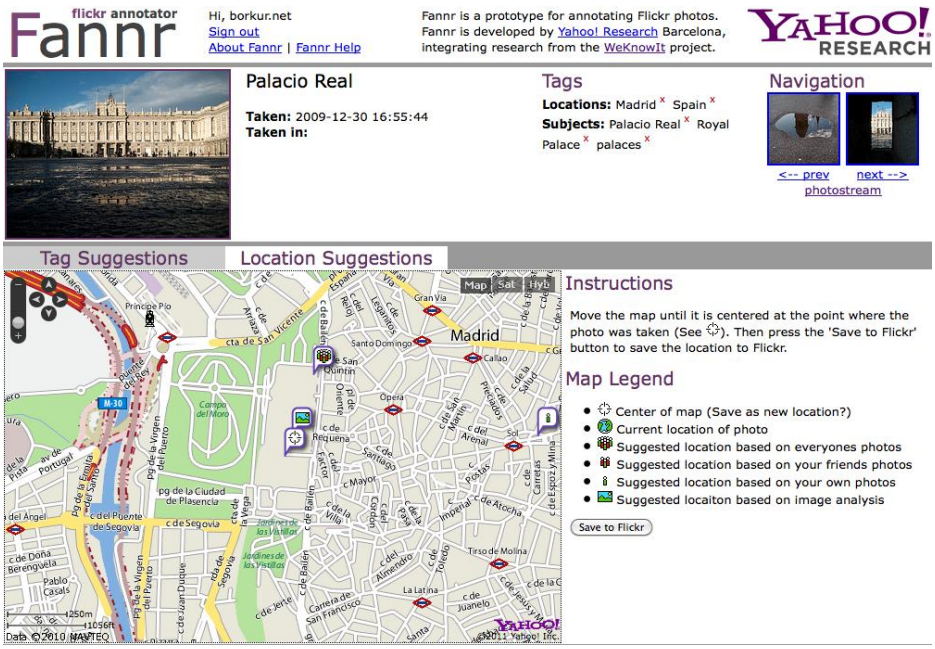

Figure 19 Fannr interface snapshot.

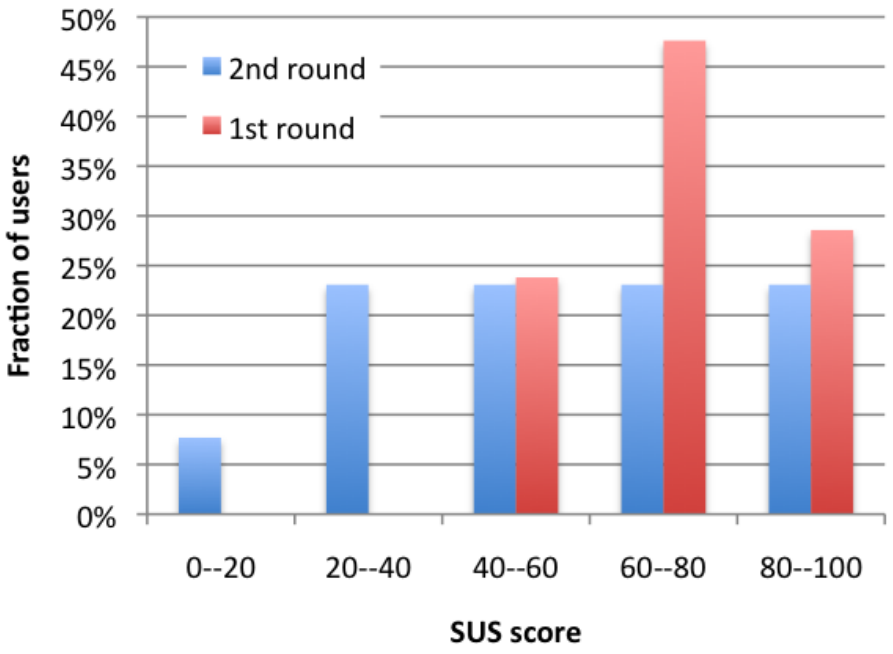

Figure 20 Travel Planner: Distribution of SUS scores.

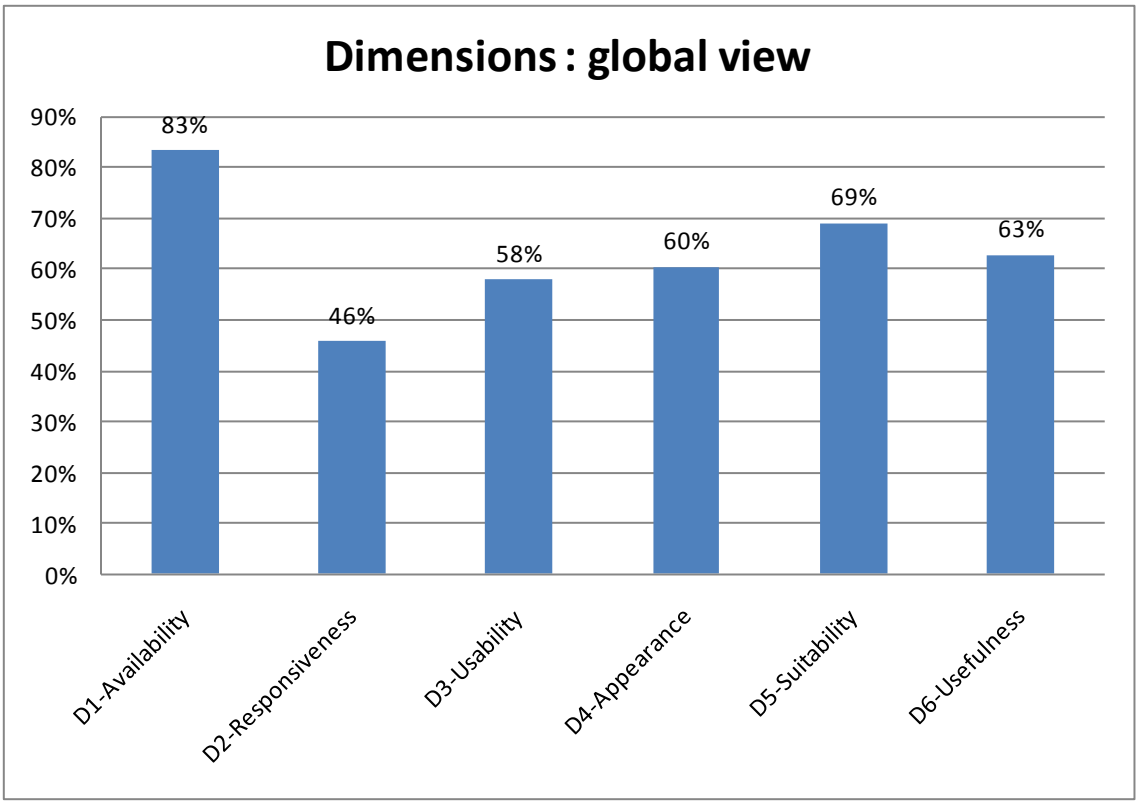

Figure 21: Mobile Guidance: Evaluation dimensions – global view.

4.1.4 Potential impact, dissemination activities and exploitation results

Impact

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.

Various domains are expected to benefit from the developed technologies in WeKnowIt, other than its two use cases (i.e. travel and tourism an emergency response). Their impact is expected to affect market, society and science. Opportunities arise in the sectors of infotainment, brand monitoring, forecast in business intelligence and analytics and security, in terms of market, as well as in cultural heritage digital libraries, e-government and environment in terms of social impact. WeKnowIt is also expected to impact scientific domains, such as computational sociology, prediction mechanisms and machine learning.

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	The unremitting servicing of end-users through the seamless

User traits	User Needs/Requirements	WeKnowIt relevance
	of any prerequisite for underlying technical knowledge	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

Market impact

The infotainment sector, and more specifically the cultural event organisation, constitutes a crucial market for European SMEs since it is highly fragmented and holds long-lasting ties with local

societies. Local European communities are proud of hosting events, both as a means of attracting tourists, and as vehicles of diffusing culture and arts through the social fabric. The capabilities that the WeKnowIt results provide can have a transformational impact on how events are experienced from the audience, and on the kind of services that are offered to event attendants. Collective Intelligence technologies for analysing massive Web 2.0 content, along with personalisation and content presentation techniques from the Personal Intelligence layer enable the possibility to make use of the project results in events of different scales, from localised ones (e.g. Thessaloniki Film Festival), to events of global reach (e.g. Cannes Festival, Olympic Games, Athens Marathon, etc.).

Brand or product monitoring are a further market where the Collective Intelligence capabilities of WeKnowIt are of great value. Such products can help professional users in the field of marketing as well as individual citizens with given information needs. The availability of new effective means for accessing massively produced web content and of new relevant applications and services will result in wider market opportunities for SMEs with content-centred and Location-Based Service related business models. The WeKnowIt technological innovations can possibly help to provide new business opportunities for SMEs in the areas of event organisation, event-related and location-based marketing and providers of information access technologies.

Another field of application for the developed WeKnowIt Collective Intelligence technologies would be the sector of business analytics and forecast. In order to improve strategic planning and increase success immediately, the tacit knowledge that lies dormant or siloed in organizations can be harnessed by tapping into the wisdom of employees, customers and partners. Using a Collective Intelligence Platform, managers can a) develop timely and more accurate business forecasts b) readily connect with employees on the ground c) align strategic decision-making d) identify and manage risks around the formulation and execution of strategic plans.

Business security can be greatly enhanced by the use of IT platforms that target crime and anti-social behaviour. Generally business protects itself by organising around two way radios, security guards, CCTV and incident logs. New Collective Intelligence platforms that can combine internet based mobile voice communications and merge it with a image distribution databases and incident management systems can the two way radio nowadays practice. Such a combination adds to the two way radio model by giving users the capability of instant, secure access to pictures, video and detailed incident management information from almost anywhere. Instead of relying solely on police led intelligence, Collective Intelligence techniques can be used for business to gather and share its own intelligence using its existing security infrastructure. This makes intelligence more relevant, local and up to date. Images from CCTV or from store operatives could be quickly uploaded to a CI system and shared with all connected businesses in a matter of seconds. Such an integration would possibly make venue security more effective and cheaper to run than the traditional approach; helping business improve its security capabilities, reduce shrinkage, make the customer experience more enjoyable, and save money on its security budget.

Social Impact

Strengthening citizens' participation in governmental and local societies decision making processes is not only a cornerstone for the notion of democracy, but also the only way for people to influence the decisions and policies that affect their daily lives. Web 2.0 based tools can allow citizens to voice their opinions, stimulate dialogue and for their views/concerns to be taken into account in governmental policies. A major drawback of current relevant applications is that they do not fully support intelligent processing and management of such user contributed information. Thus, both citizens and local decision makers fail to access the large amount of content efficiently and cannot exploit the underlying, hidden knowledge. Collective Intelligence techniques can be applied on this content to extract patterns and knowledge arising, and contribute to a policy planning which reflects better the citizens' needs. Large-scale processing which takes in account content (e.g. images, text, videos), tags, social information (e.g. friendships and citizens communities), location and spatial information can be exploited with the technologies created in WeKnowIt in order to produce Collective Intelligence, thus providing added-value to the available content and making existing procedures and workflows more efficient. For example, by jointly analysing blog comments and the location of the users submitting the comments, a specific topic (i.e. "bicycle lanes") might arise as dominant for a specific area showing the interest of the citizens on this topic and allowing efficient planning.

Moreover, analysis of user-contributed content via Collective Intelligence techniques allows linking with existing structured or semi-structured sources, e.g. by creating Linked Data. Furthermore, again in terms of Collective Intelligence technologies, Linked Data can be further combined (mashed-up) with any other piece of Linked Data, promoting the usage of the acquired data in various areas. For example, government data on health care expenditures for a given geographical area can be combined with other data about the characteristics of the population of that region in order to assess effectiveness of the government programs. A combination of web 2.0 and semantic technologies can make it possible both for citizens to track legal procedures, understand technical documents, express views, and for elected representatives to better handle the gathered information and transfer it to other bodies as well.

Specific e-Government fields that can benefit from such a Collective Intelligence framework are opinion mining aiming at topic and trends detection, public consultation procedures, or problem reporting at local (neighbourhood) level.

Nowadays cultural content and related knowledge is widely dispersed in multiple cultural institutions or on the Web. In recent years, only few institutions have managed to open their collections, to make them available in a digital form, and to provide access to these digital surrogates on the Web. Even then, data inside institutional databases often cannot easily interact with external data and usage of systems is limited by the actions offered through an interface. Often there is no possibility to link

records inside institutions which limits the value of their data. Collective Intelligence introduces advanced possibilities for cross-domain cooperation and will enable institutions – particularly from the cultural sector – to exploit new synergies, allowing their datasets being reused in completely new contexts, and enabling richer services provided for a variety of users. Collective Intelligence services can support an active role of cultural institutions to build a sound Web-based information infrastructure. Diverse, isolated data sets of cultural institutions and from the Web2.0 can be turned into linked datasets which are interlinked amongst and with existing linked datasets. As a result, cultural institution and user communities could explore and experience their multimedia resources online in an instant.

Global climate change is perhaps the most pressing and important problem currently facing humanity. It is also unique by virtue of being a truly systemic problem of vast complexity, affecting everyone and being directly affected by everyone's actions. Like nothing else, dealing with climate change calls upon us to engage in effective collective decision making on a global scale. The spectacular emergence of the Internet and associated information technology has enabled unprecedented opportunities for such interactions. To date, however, these interactions have been incoherent and dispersed, contributions vary widely in quality, and there has been no clear way to converge on well-supported decisions concerning what actions, both grand and ground-level, humanity should take to solve its most pressing problems. This important challenge could be addressed through the creation of new class of web-mediated discussion and decision making fora, that will use an innovative combination of internet-mediated interaction, collectively generated idea repositories, computer simulation, and explicit representation of argumentation to help large, diverse, and geographically-dispersed groups systematically explore, evaluate, and come to decisions concerning systemic challenges.

Scientific Impact

Computational sociology is a branch of sociology that uses computationally-intensive methods to analyze and model social phenomena. Using computer simulations, artificial intelligence, complex statistical methods, and new analytic approaches like social network analysis, computational sociology develops and tests theories of complex social processes through bottom-up modeling of social interactions. WeKnowIt Collective Intelligence techniques that enable the combination of SNA technologies with media and mass content analysis methods can boost research also in computational sociology.

An emerging new path in prediction research is to explore the global trends and sentiments that can be drawn by analyzing the sharing patterns of uploaded and downloaded social multimedia. In a sense, each time an image or video is taken or viewed, it constitutes an implicit vote for (or against) the subject of the image. This vote carries along with it a rich set of associated data including time and (often) location information. By aggregating such votes across millions of Internet users, it is

possible to reveal the wisdom that is embedded in social multimedia sites for social science applications such as politics, economics, and marketing.

The scientific and technological progress developed within WeKnowIt will have significant influence on research communities at a European and a global level in the topics investigated by the project: Data Mining and Knowledge Discovery, Multimedia Information Retrieval, Contextual Media analysis and fusion, Personalisation and Recommendation and Human Computer Interaction are the main fields that are influenced by the research work in WeKnowIt.

Dissemination activities

Consortium partners have been engaged in the pursuit of market exploitation of the project's new technologies and applications. In general, dissemination activities included publications, organisation and participation in events, the summer school and contacts with users and other projects. Exploitation activities included the creation of the WeKnowIt User Group, participation in exhibitions and industrial events, patent applications, establishment of spin-off companies and contact with companies in order to establish common exploitation strategies.

The WeKnowIt consortium has successfully participated in and organised a number of dissemination events and participated in industrial events in order to promote the project results. A summary is depicted in Table 1.

User testing has been actively pursued during the two evaluation phases of the project. Especially, with the creation of the WeKnowIt User Group, many of its members (such as Bristol City council, Manchester City Council, Niobium Labs, Sboing.net etc.) have engaged in testing the WeKnowIt results, even after the end of the project (e.g. St. Paul's Carnival in Bristol in July 2011).

Furthermore, users awareness is pursued through the issuing of the monthly WeKnowIt newsletter, presenting all the latest news and events of the project, useful information and tips. Finally the WeKnowIt video showcase has been produced and is publicly released[21].

Organization and participation in dissemination and exploitation events

During the project, the project's consortium partners have successfully participated in and organised a number of dissemination events. More specifically, VOD and CERTH-ITI participated and gave a short presentation to a workshop organised by the Technical Chamber of Greece, in Athens, on May 5, 2008, under the title: "ICT at the service of the ecosystem's and citizens' prevention and protection against forest wildfires".

[21] http://vimeo.com/weknowit

Activity	Number	Notes
Workshops	16	- TCG - ESWC ’08, CISWeb - SAMT ’08 - Sem. Web Challenge ’08 - WIAMIS ’09 - CBMI ’09 - Sem. Web Challenge ’10 - KCAP ’09, CKCaR - ACM MM ’09, EiMM - Core Cities - SETN ’10, EVENTS - ACM MM ’10, WSM - ICT Event ‘10, 1 Networking Session - ACM MM ’10, EiMM - ACM MM ’11, EiMM - ACM MM ’11, WSM
Conferences	8	- CIVR ’09 organization - WWW ’09 tutorial - IASA, project presentation - WWW ’10, 5 papers - CIVR2010 - SMIND WKI conference organization in Warsaw - ACM ’10, 2 papers, 2 demos, 1 grand challenge participation - WWW ’10, 5 papers
Publications	4	- 50 conference papers - 15 workshop papers

		- 12 journals - 2 book chapters
Media	4	- Butterfly EUROPE event - Video showcase production - Special event on WeKnowIt, Czech Republic - Video as supplementary multimedia material for the accepted article "Cluster-based Landmark and Event Detection on Tagged Photo Collections" in the IEEE Multimedia magazine.
Summer Schools	2	- SSMS ’09 - SSMS ’10
Exhibitions	6	- MWC ’10 - CeBIT ’10 - BAPCO ’10 - Infosystem ‘10 - NEM ‘10 - MWC ‘11
Clustering	6	- Emergency organizations - ESWC, “News from the front” - Pronto meetings - WeKnowIt User Group - Common exercises with emergency organizations - Cooperation with Pronto, GLOCAL and VisitoTuscany

Table 1- Dissemination and exploitation activities

WeKnowIt participated in the organization of the 1st International Workshop on Collective Semantics: Collective Intelligence & the Semantic Web (CISWeb 2008), at the European Semantic Web Conference (ESWC08) in Tenerife, Spain, by initiative of CERTH-ITI.

USFD created a website (http://nebula.dcs.shef.ac.uk/sheffieldfloods/) to collect data from citizens

about the Sheffield floods, thus advertising WeKnowIt work in the Emergency Response Scenario.

In addition, SCC has got into communications with partner organizations, with Emergency Planning Society, the Yorkshire & Humber EPO's Forum and other Local Authority Groups.

A successful (more than 100 participants) ICT 2008 networking session was organized during the Lyon ICT 2008 event in November 2008, entitled "Collective Intelligence and Social Content Semantic Analysis". The session's presentations included an Introduction to Collective Intelligence, describing the sources, processes and consumption of Collective Intelligence along with the current situation.

An invited presentation at the "Workshop on Cross-media information analysis, extraction and management at SAMT 2008" was also organized, which took place in Koblenz, Germany on 3-5 December 2008. Moreover, SemaPlorer, an application created by University of Koblenz, won the Billion Triples 1st prize in the Elsevier sponsored 2008 Semantic Web Challenge.

A WeKnowIt derived tutorial entitled "Detecting, Understanding and Exploiting Web communities" has been presented in the 18th International World Wide Web Conference.

For the Semantic Web Challenge 2010, which was held in the context of the 9th International Semantic Web Conference in Shanghai, China from November 7.-11, the SemaPlorer application was selected among the finalists.

UoKob has organised the Summer School on Multimedia Semantics, Managing and Modeling of Multimedia and User Generated Content in Web 3.0, Koblenz, Germany, 24-28 August, 2009. During this Summer School, invite talks were given by speakers on media analysis, semantic annotation and semantic retrieval. Finally, also lectures on social media semantic multimedia modelling were given. URL: http://www.smart-society.net/ssms09/ .

UoKob and CERTH-ITI also organised the 1st International Workshop on Collective Knowledge Capturing and Representation (CKCaR) in conjunction with KCAP, in California, USA, on September 2009. URL: http://www.uni-koblenz.de/confsec/CKCaR09/

UoKob were the organisers of the 1st ACM International Workshop on Events in Multimedia (EiMM), which was held in conjunction with the ACM Multimedia Conferennce, Beijing, China, October 19-24, 2009. URL: http://www.uni-koblenz.de/confsec/eimm09/

WeKnowIt was the organiser of the 10th International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS 2009), 6-8 May, London, UK. The International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS) is one of the main international fora for the presentation and discussion of the latest technological advances in interactive multimedia services. The objective of the workshop was to bring together researchers and developers from academia and

industry working in all areas of image, video and audio applications, with a special focus on analysis. URL http://wiamis2009.qmul.net/

WeKnowIt organised the ACM International Conference on Image and Video Retrieval (CIVR 09), 8-10 July, 2009, Island of Santorini, Greece. The International Conference on Image and Video Retrieval (CIVR) series of conferences was originally set up to illuminate the state of the art in image and video retrieval between researchers and practitioners throughout the world. This conference aimed to provide an international forum for the discussion of challenges in the fields of image and video retrieval. URL: http://www.civr2009.org/

WeKnowIt has organised the 7th International Workshop on Content-Based Multimedia Indexing (CBMI 09),3-5 June, 2009, Chania, Crete, Greece. CBMI 2009 brought together the various communities involved in the different aspects of content-based multimedia indexing, such as image processing and information retrieval with current industrial trends and developments. URL: http://image.ntua.gr/cbmi2009/

At a meeting of the Core Cities, which represents the 8 prominent cities in England outside London, SCC staged a workshop with Sheffield University at which the majority of Emergency Planning lead Officers took part comparing traditional reporting methods during an emergency with the potential advantages of the WeKnowIt project. As a result of the workshop considerable interest has been generated and two of the cities have indicated a desire to be further involved in the Project.

WeKnowIt participated in the annual conference of the International Association of Sound and Audiovisual Archives (IASA) that was held in Athens on 20-25 September 2009. An extended half-hour presentation of the project was given by CERTH-ITI.

WeKnowIt, in cooperation with the European ICT project Pronto , will co-organize the 1st International Workshop on recognising and tracking events on the Web and in real life (EVENTS 2010), which will be hosted by the 6th Hellenic Conference on Artificial Intelligence (SETN 2010) conference in Athens, Greece, May 4, 2010. On behalf of WeKnowIt, CERTH-ITI is taking part in the Organisational Committee. URL: http://mklab.iti.gr/events2010/

Five WeKnowIt relevant and supported papers were accepted for the World Wide Web Conference 2010.

WeKnowIt participated in bilateral meetings during a Brokerage Event held in the context of the Mobile World Congress industrial event during 15-18 February 2010 in Barcelona, with the aim to promote the project exploitation potential and achieve industrial cooperation. A number of interesting meetings took place, which resulted in the enrichment of the WeKnowIt User Group.

WeKnowIt participated in bilateral meetings during the "Future Match" event in the context of the CeBIT digital industry tradeshow from 2-6 March in Hannover, aiming at further promoting the

project exploitation plans and lead the industrial drive of the project. In these meetings very interesting contacts were established between WeKnowIt and the industry field of the WeKnowIt use cases (as well as other domains).

WeKnowIt participated in the "News from the Front" project matchmaking session during the upcoming European Semantic Web conference, which will be held from May 30 to June 3 in Heraklion, Greece. WeKnowIt had the opportunity to present itself to the wider research community and explore potential synergies and co-operations with other projects in a focused and constructive manner.

In the cooperation framework between WeKnowIt and Pronto projects, bilateral meetings were organised and hosted by members of both consortia.

SCC, together with colleagues from USFD, took part at the BAPCO 2010 conference and exhibition on 20-22 April in London, UK (http://www.bapco.org.uk/). BAPCO is well renowned amongst key industries and organisations. The conference attracted delegates from around the world. The exhibition was well received and provided useful networking opportunities. As a result of the conference, SCC and USFD were approached by a number of industry magazines to provide future WeKnowIt articles.

WeKnowIt in collaboration with Pronto co-organized the 1st International Workshop on recognising and tracking events on the Web and in real life (EVENTS 2010), which was hosted by the 6th Hellenic Conference on Artificial Intelligence (SETN 2010) conference in Athens, Greece, May 4, 2010. URL: http://mklab.iti.gr/events2010/

Also, WeKnowIt supported the organisation of the 2nd ACM Int. Workshop on Events in Multimedia (EiMM10), which was held in conjunction with the ACM Multimedia 2010 conference in Firenze, Italy, October 25-29, 2010. The 2nd Workshop on Events in Multimedia is a joined effort of the WeKnowIt project and GLOCAL project. URL: http://www.uni-koblenz.de/confsec/eimm10/

WeKnowIt supported and participated in the organization of the ACM 2nd SIGMM International Workshop on Social Media (WSM2010) in conjunction with ACM Multimedia Conference, October 25-29, 2010, Firenze, Italy. http://www.cais.ntu.edu.sg/~wsm2010/

WeKnowIt organised the ICT event 2010 networking session on Web of Events on October 27th, 2010, in Brussels. URL: http://icep-ict2010.fzi.de/

The consortium organised its participation in the NEM Summit 2010 event in Barcelona. All partners contributed by preparing live demonstrations of the WKI tools and prototypes, posters and video demos. In this context WeKnowIt produced an updated flyer22. Overall, the demo booth collected

[22] http://www.weknowit.eu/documents/weknowit_flyer

good impressions. URL: http://nem-summit.eu/

WeKnowIt also sponsored the SSMS 2010 Summer School on Multimedia Semantics, which took place during August 30 - September 3 in Amsterdam, Netherlands. The summer school was aimed at PhD students who want to learn about the connections between media analysis and knowledge representation, including basic and more advanced lectures in both fields. URL: http://ssms10.project.cwi.nl/

WeKnowIt was presented at the ACM's International Conference on Image and Video Retrieval (CIVR 2010) Practicioner's Day event on July 7th, 2010, in Xi'an, China. url: http://www.civr2010.org/programe.htm

WeKnowIt supports the organisation of the 3rd ACM Int. Workshop on Events in Multimedia (EiMM11), which will be held as part of the Joint ACM Workshop on Modeling and Representing Events (J-MRE'11), in conjunction with the ACM Multimedia 2011 conference in Scotsdale, Arizona, USA, November 28 - December 1, 2011. The 3rd Workshop on Events in Multimedia is a joined efford of the WeKnowIt project and GLOCAL project. URL: http://www.acmmm11.org/content-workshops.html

WeKnowIt supports the organization of the ACM 3rd SIGMM International Workshop on Social Media (WSM2011) in conjunction with the ACM Multimedia 2011 conference in Scottsdale, Arizona, USA, November 28 - December 1, 2011. URL: http://www.cais.ntu.edu.sg/~wsm2011/

SMIND organised a workshop which was part of a prestigious business conference held by Harvard Business Review (Warsaw, Poland). This event was aimed at CEOs and managers responsible for determining of strategic business development decisions. The presentation of ideas and technologies developed within the WeKnowIt project aimed at propagation of knowledge about semantic technologies and their uses in business.

WeKnowIt was successfully presented in the worldwide premier multimedia conference, ACM Multimedia Conference 2010 in Florence with support for 2 workshops (on events and social media), 2 oral papers (acceptance rate ~16%), 2 demo presentations and 1 Multimedia Grand Challenge participation. URL: http://www.acmmm10.org/

WeKnowIt presentation at the EU Social Networks and EU R&I Programmes Workshop, 26 Nov. 2010, Brussels. URL: http://cordis.europa.eu/fp7/ict/programme/events1-sn-2010_en.html

Participation in the 6th Networked Media Systems Unit FP7 Concertation Meeting, 29 Nov. 2010 Brussels. URL: http://www.ist-chorus.org/Cluster%20meetings

BUT organized a special national event on WeKnowIt, with the participation of mainly Czech but also Slovak audience, as part of the regular meeting which took place at the Faculty of Information

Technology, BUT, in November 2010. Pavel Smrz as the chairman presented the results of the speech processing services as well as the general context of the current project development.

A comparative evaluation exercise was held at South Yorkshire Fire and Rescue Training and Development Centre, Sheffield, in March 2011 with participation of ER staff, Officers from Manchester, Bristol & Rotherham Councils, Citizens and Sheffield College students. Its successful results will appear in the Emergency Services Times magazine (June 2011).

WeKnowIt was invited to participate at St. Paul Carnival in Bristol, on the 2nd July, 2011. The Emergency Response prototype will be part of the multi-agency tactical command Control Room, situated ~ 500m. from the festival site. 6 Smartphones with the ER application will be given to ER personnel in the streets, while photo/video/audio evidence of the set up of the carnival, the carnival procession and the street parties afterward will be collated. A laptop connected to a wide wall screen will be set up in the control room to review the incoming information and ER officers will be provided with WKI smartphones to test the system.

ClustTour was presented and demoed in the InfoSystem ICT exhibition in Thessaloniki in the booth of CERTH-ITI.

WeKnowIt participated in bilateral meetings during a Brokerage Event held in the context of the Mobile World Congress industrial event during 14-17 February 2011 in Barcelona, with the aim to promote the project exploitation potential and achieve industrial cooperation. A number of interesting meetings took place, which resulted in the enrichment of the WeKnowIt User Group. Also, possibilities for further cooperation have arisen, especially for the Emergency Response use case.

Exploitation activities and first results

With respect to exploitation activities, apart from the project participation in industrial events, as mentioned above, WeKnowIt has formed and populated its User Group. Individual partners and the Coordinator were engaged in spotting and contacting candidate organizations. At the end of the project, the list comprises the following organisations:

1. Stadt Dortmund Fire Department (FDDO), Germany
2. FKT - Baden Baden, Germany
3. Bristol City Council, UK
4. DEL REY Systems & Technology Inc., CA, USA
5. Novageo Solutions, Portugal
6. Manchester City Council, UK
7. CW Security Communications
8. Sboing.net
9. CWSEC Ltd, UK

10. EADS IW, UK
11. VELTI S.A., Greece
12. 21media innovations ltd, UK
13. Innova SpA , Italy
14. IN2 Search Interfaces Development ltd, Germany
15. VRT-medialab, Belgium
16. Somerco, UK
17. Tigatag, UK

Further, new collaborations between WeKnowIt and small IT companies have started to emerge, so as to make the project results exploitable in a broader context. Discussions between small IT companies are going on with individual partners regarding possible collaboration and exploitation of the project results.

Plans for exploiting the VIRaL image retrieval and localization tool have been set up oinvolving cooperation between CERTH and VOD.

CERTH-ITI considers ClustTour as a very promising application and technology and has thus explored multiple avenues for the exploitation of the application and/or the underlying content analysis technology. ClustTour was submitted as a business idea to the OpenFund - a Greek venture capital aiming at financing and supporting promising high-tech start ups. Although the idea was not selected for funding, it received positive feedback with respect to its technical aspects and valuable recommendations regarding the potential for exploiting the underlying technologies. Reviewers pointed the utility of the service for travelers, as well as the technical soundness of the idea and the development team. In addition, they recommended the redefinition of the business model such that the prospective business can be sustainable from well-specified (and plausible) revenues. In addition, ClustTour was among the 40 technologies (out of 150) that were selected as mature technologies for exploitation by the INTERVALUE[23] platform. The INTERVALUE Platform constitutes a repository containing research results from European Research Organisations that lead to the development of new products, new production processes and new services. Technology providers, from universities and other R&D institutions, submit information about research products and services, while technology users, from both the private and public sector, can access this information over the web. Recently, CERTH-ITI has initiated discussions with imagga about the potential of applying the image clustering technology underlying ClustTour on the StockPodium photo repository with the goal of improving the browsing and discoverability of stock photos. It has been agreed that preliminary tests will be performed on a set of photos provided by imagga and in case of satisfactory results a licensing agreement will be signed. Currently, imagga is looking into the legal aspects of making StockPodium photo content and accompanying metadata available to CERTH-ITI for

[23] http://www.researchvalue.net/

performing a series of evaluation tests. CERTH-ITI has also applied for a patent to the Hellenic Industrial Property Organisation regarding the ClustTour framework.

Furthermore, two companies are currently in touch with USFD to discuss possible commercial exploitations or proof of concepts, namely Seashore Argyll and K-Now, a spin-off company of the OAK Group (University of Sheffield). Finally, USFD has obtained a grant to follow up the work done in the Personal and Media Intelligence layers of WeKnowIt, as well as in the Emergency Response use case. The project is entitled "Tracking Real-time Intelligence in Data Streams (TRIDS)" and is funded by the UK Technology and Strategy Board SBRI programm, under the call - Have I Got 'Views' For You?[24.] In its first phase that will last from July-December 2011, the project will provide a software solution able to gather, collate and analyse in real time a large number of live data streams, linking the live data to static information sources in order to foster understanding of unfolding events. The approach is based on modelling Collective Intelligence, i.e. the combination of personal intelligence (what the single individuals communicate, e.g. tweets), mass intelligence (e.g. trends in media), social intelligence (e.g. group/community opinions) and organisational intelligence (e.g. official and professional bodies communications). Through this project the UK government will be able to monitor events via social sites, by directly exploiting results from WeKnowIt. It is highly probable that this project will continue with a second phase in 2012.

WeKnowIt research led to a commercial feature of serving ranked entity facets in Yahoo! Image Search and Yahoo! Web Search. Media Intelligence was used to extract facets from semi structured information sources. Mass Intelligence was used to rank the facets using query logs and Flickr photo annotations and to train a machine learned ranking approach using click data. The feature gained significant traction and increased number of page impressions.

With respect to speech analysis, BUT has provided new enhancements in speech recognition technology applied to Phonexia Ltd. Services. BUT has also established links with OptimSys Ltd. (VoiceXML solutions) & Lingea Ltd. (dictionary s/w vendor) to offer ASR solutions developed in WeKnowIt. The WeKnowIt-based speech indexing & recognition services have been integrated in Přednášky (a lecture portal).

EMKA plans to launched to market a developed mobile app for emergency alert situations with the cooperation of 3 German companies (ASE AG, Nettropolis AG, RA Consulting GmbH).

A great amount of patents have been filed for by members of the consortium. Specifically, Yahoo! has filed for seven patents, while CERTH-ITI and TID have filed for one patent each.

Finally, UoKob has started a spin-off company Kreuzverweis, exploiting results of the Organisational Intelligence layer. The company's goal is to alleviate today's challenges in media

[24] http://www.innovateuk.org/content/competition/have-i-got-views-for-you-.ashx

management regarding the creation, indexing, and usage of media content. The Kreuzverweis Solutions GmbH has been founded in January 2011. It is supported by the research transfer programme EXIST of the Federal Ministry of Economics and Technology, Germany. A first product of Kreuzverweis has been released in February 2011. The Kreuzverweis Annotator is a plugin for Aperture, a commercial photo management software for Macintosh.

4.1.5 Contacts and Information

For further information visit the project web site http://www.weknowit.eu or send e-mail to the project coordinator Dr.Yiannis Kompatsiaris, ikom@iti.gr.

4.2 Use and dissemination of foreground

Section A (public)

TEMPLATE A1: LIST OF SCIENTIFIC (PEER REVIEWED) PUBLICATIONS, STARTING WITH THE MOST IMPORTANT ONES										
NO.	Title	Main author	Title of the periodical or the series	Number, date or frequency	Publisher	Place of publication	Year of publication	Relevant pages	Permanent identifiers[25] (if available)	Is/Will open access[26] provided to this publication?
1	The Dynamics of Content Popularity in Social Media	Symeon Papadopoulos	International Journal of Data Warehousing and Mining	Vol. 6(1)	IGI Global	USA	2010	pp. 20-37	http://www.igi-global.com/bookstore/article.aspx?titleid=38952	No
2	SemaPlorer—Interactive Semantic Exploration of Data and Media based on a Federated Cloud Infrastructure	S. Schenk	Journal of Web Semantics	7(4)	Elsevier	Netherlands	2009	pp. 298-304	10.1016/j.websem.2009.09.006	Yes
3	An Ecosystem for Semantics	A. Scherp	IEEE MultiMedia	16(2)	IEEE	USA	2009	pp.18-	http://ieeexplor	No

[25] A permanent identifier should be a persistent link to the published version full text if open access or abstract if article is pay per view) or to the final manuscript accepted for publication (link to article in repository).

[26] Open Access is defined as free of charge access for anyone via Internet. Please answer "yes" if the open access to the publication is already established and also if the embargo period for open access is not yet over but you intend to establish open access afterwards.

								25	e.ieee.org/xpls/abs_all.jsp?arnumber=5167286&tag=1	
4	An update algorithm for restricted random walk clustering for dynamic data sets	Markus Franke	Advances in Data Analysis and Classification	3(1)	Springer	UK	2010	pp. 63-92	http://www.springerlink.com/content/b535555613325431/	No
5	Cluster-based Landmark and Event Detection on Tagged Photo Collections	Symeon Papadopoulos	IEEE Multimedia Magazine	18(1)	IEEE	USA	2011	pp. 52-63	http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5611558	No
6	Using Visual Context and Region Semantics for High-Level Concept Detection, IEEE Transactions on Multimedia.	Phivos Mylonas	IEEE Transactions on Multimedia	11(2)	IEEE	USA	2009	pp. 229-243	http://portal.acm.org/citation.cfm?id=1652981	No
7	Concept Detection and Keyframe Extraction Using a Visual Thesaurus, Multimedia Tools and Application.	Evaggelos Spyrou	Multimedia Tools and Applications	41(3)	Kluwer Academic Publishers	Hingham, MA, USA	2009	pp. 337-373	http://portal.acm.org/citation.cfm?id=1499074	No
8	Harvesting Intelligence in Multimedia Social Tagging Systems, Emergent Web	Eirini Giannakidou	Emergent Web Intelligence, Springer Verlag, Series: Studies in Computational		Springer	UK	2010	pp. 135-167	http://dx.doi.org/10.1007/978-1-84996-074-	No

	Intelligence, Springer Verlag, Series: Studies in Computational Intelligence		Intelligence						8_6	
9	Approaches to Visualising Linked Data: A Survey, (Accepted for publication in) the Semantic Web Journal — Special Call for Survey articles on Semantic Web topics.	Dadzie, A.-S	Semantic Web Journal	To appear	IOS Press	Netherlands	2011		http://www.semantic-web-journal.net/content/new-submission-approaches-visualising-linked-data-survey	No
10	Community Detection in Social Media	Symeon Papadopoulos	Data Mining and Knowledge Discovery journal	To appear	Springer	UK	2011			No
11	A Core Ontology on Events for the Semantic Representation of Human Experiences in the Real World	A. Scherp	Multimedia Tools and Applications	To appear	Springer	Netherlands	2011		http://www.springerlink.com/content/72013082v64lv5n5/	No
12	VIRaL: Visual Image Retrieval and Localization	Y. Kalantidis	Multimedia Tools and Applications	51(2)	Springer	Netherlands	2011	pp. 555-592	10.1007/s11042-010-0651-7	No
13	Web 2.0 and Traditional Knowledge Management Processes	A. Scherp	KSM09 - 1st Workshop on Knowledge Services & Mashups	March 2009		Solothurn, Switzeralnd	2009	pp. 222-231	http://subs.emis.de/LNI/Proceedings/Procee	Yes

									dings145/article2416.html	
14	Lexical Graphs for Improved Contextual Ad Recommendation	Symeon Papadopoulos	Lecture Notes in Computer Science,	Vol. 5478/2009	Springer	UK	2009	pp. 216-227	10.1007/978-3-642-00958-7_21	No
15	User Requirements for a Collective Intelligence Emergency Response System	V. Lanfranchi	The 23rd BCS conference on Human Computer Interaction	September 2009	British Computer Society	Cambridge, UK	2009	pp. 198-203	http://portal.acm.org/citation.cfm?id=1671035	No
16	Local Community Situational Awareness During An Emergency	N. Ireson	3rd IEEE International Conference on Digital Ecosystems and Technologies	June 2009		Istanbul, Turkey	2009	pp.49-54	http://ieeexplore.ieee.org/xpls/abs_all.jsp?tp=&arnumber=5276763	No
17	Leveraging Web 2.0 Communities in Professional Organisations	A. Scherp	W3C Workshop on the Future of Social Networking	January 2009		Barcelona, Spain	2009		http://mklab.iti.gr/content/leveraging-web-20-communities-professional-organisations	No
18	Large Scale Concept Detection in Video Using a Region Thesaurus	E. Spyrou	Lecture Notes in Computer Science	Volume 5371/2009	Springer	UK	2009	pp. 197-207	10.1007/978-3-540-92892-8_20	No
19	Visual Image Retrieval and Localization	Y. Kalantidis	7th International Workshop on Content-Based	June 2009		Chania, Greece	2009		http://www.image.ntua.gr/php/	No

			Multimedia Indexing						pub_details.php?code=588	
20	Collective Intelligence Emergency Response System	V. Lanfranchi	International Workshop on Mobile Information Technology for Emergency Response	May 2009		Gothenburg, Sweden	2009			No
21	Placing Flickr photos on a Map	P. Serdyukov	Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval	July 2009	ACM New York	NY, USA	2009	pp. 484-491	10.1145/1639714.1639744	No
22	Ozone Browser: Augmenting the Web with Semantic Overlays	G. Burel	Semantic Web Scripting Challenge, ESWC 2009	Vol. 449/2009		Heraklion, Greece	2009		http://CEUR-WS.org/Vol-449/Challenge1.pdf	Yes
23	Leveraging Collective Intelligence through Community Detection in Tag Networks	Symeon Papadopoulos	First International Workshop on Collective Knowledge Capturing and Representation - CKCaR'09	September 2009		Redondo Beach, CA, USA	2009		http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.178.3179	No
24	Investigation into bottle-neck features for meeting speech recognition	F. Grezl	InterSpeech 2009	September 2009	ISCA	Brighton, UK	2009	pp. 2947-2950	http://www.isca-speech.org/archive/interspeech_2009/i09_2947.html	Yes

25	Rate it Again: Increasing Recommendation Accuracy by User re-Rating	X. Amatriain	Proceedings of the third ACM conference on Recommender systems	October 2009	ACM New York	New York, NY, USA	2009	pp. 173-180	http://portal.acm.org/citation.cfm?doid=1639714.1639744	No
26	F-A Model of Events based on the Foundational Ontology DOLCE+ Ultra Light	A. Scherp	Proceedings of the fifth international conference on Knowledge capture	September 2009	ACM New York	Redondo Beach, CA, USA	2009	pp. 137-144	http://portal.acm.org/citation.cfm?id=1597760	No
27	Interaction and User Experiences with Multimedia Technologies---Challenges and Future Topics	A. Scherp	Proceeding of the 3rd ACM international workshop on Human-centered computing	October 2008	ACM New York	Vancouver, Canada	2008	pp. 1-6	http://portal.acm.org/citation.cfm?doid=1462027.1462028	No
28	GeoFolk: Latent Spatial Semantics in Web 2.0 Social Media	S. Sizov	Proceedings of the third ACM international conference on Web search and data mining	February 2010	ACM New York	New York, NY, USA	2010	pp. 281-290	http://portal.acm.org/citation.cfm?doid=1718487.1718522	No
29	A region-based visual vocabulary for efficient image retrieval	A. Scherp	International Workshop on Image Analysis for Multimedia Interactive Services	April 2010		Desenzano del Garda, Italy	2010			No
30	Sub-linear Indexing of Appearance and Global Geometry	Y. Avrithis	IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2010)	June 2010		San Francisco, CA, USA	2010			No
31	Intelligent Content Retrieval using a Visual Vocabulary	E. Spyrou	IEEE International Conference on Fuzzy	July 2010	IEEE	Barcelona, Spain	2010	pp. 1-8	10.1109/FUZZY.2010.55840	No

	and Geometric Constraints		Systems (FUZZ), 2010						00	
32	Modeling and Proving Trustworthiness of Web Resources using Veracity	G. Burel	ESWC 2010	June 2010		Heraklion, Greece	2010		http://oro.open.ac.uk/id/eprint/26685	No
33	Attention-Streams: Attention Based Real-time Recommendations	G. Burel	ESWC 2010	June 2010		Heraklion, Greece	2010			No
34	Collaborative Semantic Points of Interest. Demo	M. Braun	ESWC 2010, Lecture Notes in Computer Science	Vol. 6089/2010	Springer	Heraklion, Greece	2010	pp. 365-369	http://www.springerlink.com/content/2g12pg6q583882w7/	No
35	Representing Distributed Groups with dgFOAF	F. Schwagereit	ESWC 2010, Lecture Notes in Computer Science	Vol. 6089/2010	Springer	Heraklion, Greece	2010	pp. 181-195	http://www.springerlink.com/content/00u168004937x1j2/	No
36	Unlocking the Semantics of Multimedia Presentations in the Web with the Multimedia Metadata Ontology	C. Saathoff	Proceedings of the 19th international conference on World wide web	April 2010	ACM New York	Raleigh, NC, USA	2010	pp. 831-840	http://portal.acm.org/citation.cfm?doid=1772690.1772775	No
37	Collective Intelligence in Mobile Consumer Social Applications	S. Diplaris	Ninth International Conference on Mobile Business and Ninth Global Mobility Roundtable (ICMB-GMR),	June 2010		Athens, Greece	2010	pp. 206 - 212	10.1109/ICMB-GMR.2010.71	No
38	Ranking Entity Facets based on User Click Feedback	R. van Zwol	33rd Annual ACM SIGIR Conference	July 2010	ACM	Geneva, Switzerland	2010			No

39	Ranking Entity Facets based on User Click Feedback	R. van Zwol	IEEE Fourth International Conference on Semantic Computing (ICSC)	September 2010	IEEE	Pittsburgh, PA, USA	2010	pp. 192-199	10.1109/ICSC.2010.33	No
40	Knowledge Sharing in eCollaboration	N. Ireson	EGOV'10 Proceedings of the 9th IFIP WG 8.5 international conference on Electronic government	September 2010	Springer	Lausanne, Switzerland	2010	pp. 351-362	http://portal.acm.org/citation.cfm?id=1887168	No
41	Image Clustering through Community Detection on Hybrid Image Similarity Graphs	Symeon Papadopoulos	17th IEEE International Conference on Image Processing (ICIP)	September 2010	IEEE	Hong, Kong, China	2010	pp. 2353 - 2356	10.1109/ICIP.2010.5653478	No
42	A Graph-based Clustering Scheme for Identifying Related Tags in Folksonomies	Symeon Papadopoulos	DaWaK'10 Proceedings of the 12th international conference on Data warehousing and knowledge discovery	September 2010	Springer	Bilbao, Spain	2010	pp. 65-76	http://portal.acm.org/citation.cfm?id=1881931	No
43	Getting Help In A Crowd - A Social Emergency Alert Service	A. Geyer-Schulz	ICE-B'2010	July 2010	Springer	Athens, Greece	2010	pp. 1-12	http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5740440	No
44	Breaking the Bystander Effect using Mobile Devices and Social Networks	A.C. Sonnenbichler	International Network for Social Network Analysis (INSNA), 30th SUNBELT	July 2010		Riva del Garda, Italy	2010			No

			Conference							
45	A Social Emergency Alert Service - A Location-Based Privacy-Aware Personal Safety Service	M. Ovelgönne	Fourth International Conference on Next Generation Mobile Applications, Services and Technologies (NGMAST)	July 2010		Amman, Jordan	2010	pp. 84 - 89	10.1109/NGMAST.2010.27	No
46	On the Hierarchicalness of Q&A Posting Networks	M. Ovelgönne	Proceedings of the 16th ACM international conference on Supporting group work	November 2010	ACM New York	Sanibel, FL, USA	2010	pp. 117-120	http://portal.acm.org/citation.cfm?doid=1880071.1880090	No
47	Cluster Cores and Modularity Maximization	M. Ovelgönne	IEEE International Conference on Data Mining Workshops	December 2010	IEEE	Sydney, Australia	2010	pp.1204-1213	http://www.computer.org/portal/web/csdl/doi/10.1109/ICDMW.2010.63	No
48	Network Growth and the Spectral Evolution Model	J. Kunegis	Proceedings of the 19th ACM international conference on Information and knowledge management		ACM New York	Toronto, Canada	2010	pp. 739-748	http://portal.acm.org/citation.cfm?id=1871437&picked=prox&CFID=25561773&CFTOKEN=11427452	No
49	Representing, Proving and Sharing Trustworthiness of Web Resources Using	G. Burel	EKAW'10 Proceedings of the 17th international conference on Knowledge	October 2010	Springer-Verlag Berlin	Lisbon, Portugal	2010	pp. 421-430	http://portal.acm.org/citation.cfm?id=19483	No

	Veracity		engineering and management by the masses						36	
50	Retrieving Landmark and Non-Landmark Images from Community Photo Collections	Y. Avrithis	Proceedings of the international conference on Multimedia	October 2010	ACM New York	Firenze, Italy	2010	pp. 153-162	http://portal.acm.org/citation.cfm?doid=1873951.1873973	No
51	ClustTour: City Exploration by use of Hybrid Photo Clustering	Symeon Papadopoulos	Proceedings of the international conference on Multimedia	October 2010	ACM New York	Firenze, Italy	2010	pp. 1617-1620	http://portal.acm.org/citation.cfm?doid=1873951.1874302	No
52	Feature Map Hashing: Sub-linear Indexing of Appearanceand Global Geometry	Y. Avrithis	Proceedings of the international conference on Multimedia	October 2010	ACM New York	Firenze, Italy	2010	pp. 231-240	http://portal.acm.org/citation.cfm?doid=1873951.1873985	No
53	Prediction of Favourite Photos using Textual, Visual, and Social Signals	R. van Zwol	Proceedings of the international conference on Multimedia	October 2010	ACM New York	Firenze, Italy	2010	pp. 1015-1018	http://portal.acm.org/citation.cfm?doid=1873951.1874138	No
54	STEVIE---Collaborative Creation and Exchange of Events and POIs on a Mobile Phone	M. Braun	Proceedings of the 2nd ACM international workshop on Events in multimedia	October 2010	ACM New York	Firenze, Italy	2010	pp. 35-40	http://portal.acm.org/citation.cfm?doid=1877937.1877948	No
55	Integrated Mobile Visualization and Interaction of Events and POIs	D. Schmeish	Proceedings of the international conference on Multimedia	October 2010	ACM New York	Firenze, Italy	2010	pp. 1567-1570	http://portal.acm.org/citation.cfm?doid=187	No

									3951.1874288	
56	Using Event Representation and Semantic Enrichment for Managing and Reviewing Emergency Incident Logs	Symeon Papadopoulos	Proceedings of the 2nd ACM international workshop on Events in multimedia	October 2010	ACM New York	Firenze, Italy	2010	pp. 41-46	http://portal.acm.org/citation.cfm?doid=1877937.1877950	No
57	Toponym Resolution in Social Media	N. Ireson	ISWC'10 Proceedings of the 9th international semantic web conference on The semantic web	November 2010	Springer-Verlag Berlin	Shangai, China	2010	pp. 370-385	http://portal.acm.org/citation.cfm?id=1940306	No
58	Using a Region and Visual Word Approach Towards Semantic Image Retrieval	Y. Kalantidis	SMAP 2010: 5th International Workshop on Semantic Media Adaptation and Personalization	December 2010		Limassol, Cyprus	2010	pp. 85-89	http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5706869	No
59	Off the beaten track - a mobile field study exploring the long tail of tourist recommendations	N. Tintarev	Proceedings of the 12th international conference on Human computer interaction with mobile devices and services	September 2010	ACM New York	Lisbon, Portugal	2010	pp. 209-218	http://portal.acm.org/citation.cfm?doid=1851600.1851636	No
60	Tú, Yo, Nosotros viajamos: La inteligencia colectiva al servicio del viajero	AF Cuadrado	XX Jornadas de Telecom I+D conjuntas con IX Jornadas de Ingeniería Telemática JITEL 2010	September 2010		Valladolid, Spain	2010		http://telecomid2010.tid.es/ranking	No
61	Faceted exploration of image	R. van Zwol	Proceedings of the 19th	April 2010	ACM New	Raleigh,	2010	pp. 961-	http://portal.ac	No

	search results		international conference on World wide web		York	NC, USA		970	m.org/citation.cfm?doid=1772690.1772788	
62	TagExplorer: Faceted Browsing of Flickr Photos	B. Sigurbjörnsson	Yahoo! Labs Technical Report (YL-2010-005)		Yahoo! Labs	USA	2010		http://vitae.borkur.net/2010/08/tagexplorer-faceted-browsing-of-flickr-photos.html	Yes
63	Who uses web search for what: and how	I. Weber	Proceedings of the fourth ACM international conference on Web search and data mining	February 2011	ACM New York	Hong Kong, China	2011		http://portal.acm.org/citation.cfm?doid=1935826.1935839	No
64	Emerging, Collective Intelligence for personal, social and organisational use	Sotiris Diplaris	"Next Generation Data Technologies for Collective Computational Intelligence", in Studies in Computational Intelligence book series	To appear in Vol. 352	Springer	UK	2011			No
65	Leveraging Massive User Contributions for Knowledge Extraction	Spiros Nikolopoulos	"Next Generation Data Technologies for Collective Computational Intelligence", in "Studies in Computational Intelligence" book series	To appear in Vol. 352	Springer	UK	2011			No
66	Digg it Up! Analyzing	Symeon	ECAI 2008 Workshop on	July 2008		Patras,	2008			No

	Popularity Evolution in a Web 2.0 Setting	Papadopoulos	Mining Social Data (MSoDa 2008)			Greece				
67	Generating Collective Intelligence	Vassilios Solachidis	32nd Annual Conference - Advances in Data Analysis, Data Handling and Business Intelligence	July 2008		Hamburg, Germany	2008			No
68	Efficient Media Exploitation towards Collective Intelligence	Phivos Mylonas	32nd Annual Conference - Advances in Data Analysis, Data Handling and Business Intelligence	July 2008		Hamburg, Germany	2008			No
69	The Potential of Social Intelligence for Collective Intelligence	A. Geyer-Schulz	32nd Annual Conference - Advances in Data Analysis, Data Handling and Business Intelligence	July 2008		Hamburg, Germany	2008			No
70	Visual Context Ontology for Multimedia High-Level Concept Detection	E. Spyrou	5th International Workshop on Modeling and Reasoning in Context (MRC 2008)	June 2008		Delft, Netherlands	2008		http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.148.7864	No
71	SEMSOC: SEMantic, SOcial and Content-based Clustering in Multimedia Collaborative Tagging Systems	Eirini Giannakidou	Second IEEE International Conference on Semantic Computing (ICSC 2008)	August 2008	IEEE	Santa Clara, CA, USA	2008	pp. 128-135	http://portal.acm.org/citation.cfm?id=1446362	No
72	Using Region Semantics And Visual Context For	E. Spyrou	1st ICIP Workshop on Multimedia Information	October 2008	IEEE	San Diego, CA, USA	2008	pp. 53-56	http://ieeexplore.ieee.org/xpls	No

	Scene Classification		Retrieval: New Trends and Challenges						/abs_all.jsp?arnumber=4711689&tag=1	
73	SemaPlorer—Interactive Semantic Exploration of Data and Media based on a Federated Cloud Infrastructure	Simon Schenk	Billion Triple Challenge, Semantic Web Conference, Karlsruhe, Germany, 2008	October 2008	Elsevier	Karlsruhe, Germany	2008	pp. 298-304	http://portal.acm.org/citation.cfm?id=1660448	No
74	Who's Who-- A Linked Data Visualisation Tool for Mobile Environments,	Cano, A.E.,	Proceedings (Demo Track), the Extended Semantic Web Conference (ESWC 2011)	To appear June 2011		Heraklion, Greece	2011			No
75	(2011). Does Size Matter? When Small is Good Enough,	Gentile, A.L.	ESWC 2011 Workshop on Making Sense of Microposts (#MSM2011)	To appear June 2011		Heraklion, Greece	2011			No
76	(2011). Sensing Presence (PreSense) Ontology — User Modelling in the Semantic Sensor Web,	Cano, A.E.	ESWC 2011 Workshop on User Profile Data on the Social Semantic Web (UWeb 2011)	To appear June 2011		Heraklion, Greece	2011			No
77	(2011). Hide the Stack: Toward Usable Linked Data,	Dadzie, A.-S	Proceedings of the Extended Semantic Web Conference (ESWC 2011)	Volume I	Springer	Heraklion, Greece	2011	pp. 93-107		No
78	A Social Location-Based Emergency Service to Eliminate the Bystander Effect	A. Geyer-Schulz	Lecture Notes in Computer Science, e-Business and Telecommunications	To appear	Springer	Germany	2011			No

79	A Comparison of Agglomerative Hierarchical Algorithms for Modularity Clustering	M. Ovelgönne	Proceedings of the 35th Conference of the German Classification Society	To appear, September 2011		Frankfurt, Germany	2011			No

TEMPLATE A2: LIST OF DISSEMINATION ACTIVITIES								
NO.	Type of activities[27]	Main leader	Title	Date	Place	Type of audience[28]	Size of audience	Countries addressed
1	Presentation	VOD	"ICT at the service of the ecosystem's and citizens' prevention and protection against forest wildfires" , Technical Chamber of Greece workshop	5 May 2008	Athens, Greece	Industry, Civil Society	50	Greece
2	Workshop	CERTH	1st International Workshop on Collective Semantics: Collective Intelligence & the Semantic Web (CISWeb 2008)	2 June 2008	Tenerife, Spain	Scientific Community	50	International
3	web	USFD	Collect data from citizens about the Sheffield floods, http://nebula.dcs.shef.ac.uk/sheffieldfloods	June 2008	Sheffield, UK	Civil Society		UK

[27] A drop down list allows choosing the dissemination activity: publications, conferences, workshops, web, press releases, flyers, articles published in the popular press, videos, media briefings, presentations, exhibitions, thesis, interviews, films, TV clips, posters, Other.

[28] A drop down list allows choosing the type of public: Scientific Community (higher education, Research), Industry, Civil Society, Policy makers, Medias ('multiple choices' is possible.

4	Workshop	CERTH	Collective Intelligence and Social Content Semantic Analysis, networking session in ICT 2008	27 September 2008	Lyon, France	Scientific Community, Industry	100	EU
5	Presentation	CERTH	"Extracting Collective Intelligence from Social Content", workshop on Cross-media information analysis, extraction and management at SAMT 2008	December 2008	Koblenz, Germany	Scientific Community	50	International
6	Other	UoKob	Summer School on Multimedia Semantics, Managing and Modeling of Multimedia and User Generated Content in Web 3.0	24-28 August 2009	Koblenz, Germany	Scientific Community	50	International
7	Presentation	CERTH	"Detecting, Understanding and Exploiting Web communities", Tutorial in the 18th International World Wide Web Conference.	20 April 2009	Madrid, Spain	Scientific Community	400	International
8	Workshop	UoKob - CERTH	1st International Workshop on Collective Knowledge Capturing and Representation (CKCaR) in conjunction with KCAP	September 2009	California, USA	Scientific Community	50	International
9	Workshop	UoKob	1st ACM International Workshop on Events in Multimedia (EiMM)	October 19-24, 2009	Beijing, China	Scientific Community	50	International
10	Workshop	CERTH	10th International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS 2009)	6-8 May, 2009	London, UK	Scientific Community	50	International
11	Conference	CERTH	ACM International Conference on Image and Video Retrieval (CIVR 09)	8-10 July, 2009	Santorini, Greece	Scientific Community	300	International
12	Workshop	CERTH	7th International Workshop on Content-Based Multimedia Indexing (CBMI 09)	3-5 June, 2009	Chania, Crete	Scientific Community	150	International
13	Workshop	SCC	UK Core Cities meeting participating Emergency Planning lead Officers	June 2009	London, UK	Policy makers	20	UK

14	Presentation	CERTH	In annual conference of the International Association of Sound and Audiovisual Archives (IASA)	20-25 September 2009	Athens, Greece	Scientific Community, Industry, Policy Makers	200	International
15	Workshop	CERTH	1st International Workshop on recognising and tracking events on the Web and in real life (EVENTS 2010)	May 4, 2010	Athens, Greece	Scientific Community	30	International
16	Other	CERTH	Bilateral meetings with industry, Brokerage Event held in the context of the Mobile World Congress industrial event	15-18 February 2010	Barcelona, Spain	Industry	20	International
17	Other	UoKob	bilateral meetings with industry in "Future Match" event in the context of the CeBIT digital industry tradeshow	2-6 March, 2010	Hannover, Germany	Industry	20	International
18	Other	UoKob	“News from the Front” project matchmaking session in European Semantic Web conference	June 3, 2010	Heraklion, Greece	Industry, Scientific Community	20	International
19	Exhibition	SCC	Stand in British Association of Public Safety Communications Officers (BAPCO) 11th annual international conference and exhibition	20-22 April 2010	London, UK	Policy makers, Civil Society	150	International
20	Workshop	CERTH-Yahoo!	Second SIGMM Workshop on Social Media (WSM 2010),	October 25-29, 2010	Firenze, Italy	Scientific Community	50	International
21	Other	UoKob - CERTH	5th Summer School on Multimedia Semantics (SSMS 2010),	30 August – 3 September, 2010	Amsterdam, the Netherlands	Scientific Community	50	International
22	Other	CERTH	WeKnowIt User Group	2009-2011	N/A	Industry	20	International
23	Workshop	UoKob	2nd ACM Int. Workshop on Events in Multimedia (EiMM10)	October 25-29, 2010	Firenze, Italy	Scientific Community	50	International
24	Workshop	UoKob	ICT event 2010 networking session on Web of Events	October 27th, 2010	Brussels, Belgium	Scientific Community, Industry	100	EU

25	Exhibition	CERTH	Stand in NEM Summit 2010 exhibition	13-15 October 2010	Barcelona, Spain	Scientific Community, Industry	300	International
26	Flyer	CERTH	WeKnowIt flyer	October 2010	N/A	Scientific Community, Industry, Policy makers, Civil society	400	International
27	Presentation	CERTH	ACM International Conference on Image and Video Retrieval (CIVR 2010) Practicioner's Day	July 7th, 2010	Xi'an, China	Scientific Community, Industry	200	International
28	Workshop	UoKob	3rd ACM Int. Workshop on Events in Multimedia (EiMM11)	November 28 - December 1, 2011	Scotsdale, Arizona, USA	Scientific Community	50	International
29	Workshop	SMIND	WeKnowIt workshop at business conference held by Harvard Business Review	September 2010	Warsaw, Poland	Policy makers, Industry	50	Poland
30	Conference	CERTH, UoKobm Yahoo!	ACM Multimedia Conference: 2 workshops (on events and social media), 2 oral papers (acceptance rate ~16%), 2 demo presentations and 1 Multimedia Grand Challenge participation	October 25-29, 2010	Firenze, Italy	Scientific Community	400	International
31	Presentation	CERTH	EU Social Networks and EU R&I Programmes Workshop	26 Nov. 2010	Brussels, Belgium	Scientific Community, Industry	50	EU
32	Other	CERTH	6th Networked Media Systems Unit FP7 Concertation Meeting	29 Nov. 2010	Brussels, Belgium	Scientific Community,	50	EU

						Industry		
33	Workshop	BUT	Event on WeKnowIt at the Faculty of Information Technology	November 2010	Brno, Czech Republic	Scientific Community, Industry	50	Czech Republic, Slovakia
34	Exhibition	CERTH	Stand at InfoSystem ICT exhibition		Thessaloniki, Greece			
35	Other	CERTH	Clusttour among 40 technologies out of 150 selected as mature for valorisation by the INTERVALUE platform	December 2010	N/A	Industry	100	Greece
36	Other	CERTH	Bilateral meetings with industry, Brokerage Event held in the context of the Mobile World Congress industrial event	14-17 February 2011	Barcelona, Spain	Industry	30	International
37	Web	CERTH	WeKnowIt website: http://www.weknowit.eu	Since 2008	N/A	Scientific Community, Industry, Civil Society, Policy makers, Medias	1500+	International
38	Press release	CERTH	WeKnowIt project initiation press release	April 2008	N/A	Medias	N/A	International
39	Interview	CERTH	Coordinator interview for Butterfly Europe event	January 2011	N/A	Medias, Policy makers	150	EU
40	Interview	CERTH	Technical Coordinator video interview for Butterfly Europe event	January 2011	Brussels, Belgium	Medias, Policy makers	150	EU
41	Video	USFD	WeKnowIt ER exercise in Sheffield - on Youtube	15 March 2011	N/A	Medias, Civil Society, Policy makers	205+	International

42	Video	CERTH	WeKnowIt presence in Vimeo: WeKnowIt dissemination video (2 versions and videoclip)	30 March 2011	N/A	Medias, Civil Society, Policy makers	200+	International
43	Presentation	CERTH	WKI public presentation in website	August 2009	N/A	Medias, Civil Society, Policy makers	1500+	
44	Other	CERTH	5 Newsletters	October 2010 – May 2011	N/A	Industry	1500+	International
45	Articles published in the popular press	CERTH	“Online travel guide”, Aggelioforos newspaper article on ClustTour	15 May 2011	Thessaloniki, Greece	Medias, Civil Society	175000	Greece
46	Video	CERTH	Video on WKI cluster-based event and landmark detection - published by IEEE “Computing Now” magazine	January 2011	N/A	Scientific Community, Medias	Millions	International

Section B

Part B2

Type of Exploitable Foreground[29]	Description of exploitable foreground	Confidential Click on YES/NO	Foreseen embargo date dd/mm/yyyy	Exploitable product(s) or measure(s)	Sector(s) of application[30]	Timetable, commercial or any other use	Patents or other IPR exploitation (licences)	Owner & Other Beneficiary(s) involved
General advancement of knowledge	Information Management through Semantically Aware Overlays	NO		Methodology, Design, Architecture, Software	Information service activities	Public disclosure		USFD
General advancement of knowledge	User Modelling for Contextual Semantic Interaction	NO		Methodology, Software	Information service activities	Public disclosure		USFD
General advancement of	Photo search and	NO		Method, Design, Content, Software,	Information service activities	Public disclosure		CERTH

[19] A drop down list allows choosing the type of foreground: General advancement of knowledge, Commercial exploitation of R&D results, Exploitation of R&D results via standards, exploitation of results through EU policies, exploitation of results through (social) innovation.

[30] A drop down list allows choosing the type sector (NACE nomenclature) : http://ec.europa.eu/competition/mergers/cases/index/nace_all.html

Type of Exploitable Foreground[29]	Description of exploitable foreground	Confidential Click on YES/NO	Foreseen embargo date dd/mm/yyyy	Exploitable product(s) or measure(s)	Sector(s) of application[30]	Timetable, commercial or any other use	Patents or other IPR exploitation (licences)	Owner & Other Beneficiary(s) involved
knowledge	localization			Algorithm				
General advancement of knowledge	Tag normalization method	NO		Method / Software	Information service activities	Public disclosure		CERTH
Commercial exploitation of R&D results	VIRaL application	NO		Method / Design / Content / Software / Algorithm / Patent	Information service activities	2012	A patent is planned for 2011	CERTH
General advancement of knowledge	Spatio-temporal annotation of social texts	NO		Methodology, Software	Information service activities	Public disclosure		USFD
General advancement of knowledge	Combining and disambiguating social and organisational metadata for collective intelligence	NO		Methodology, Software	Information service activities	Public disclosure		USFD
General advancement of knowledge	Fusion of geo-coding resources	NO		Methodology, Software	Information service activities	Public disclosure		USFD

Type of Exploitable Foreground[29]	Description of exploitable foreground	Confidential Click on YES/NO	Foreseen embargo date dd/mm/yyyy	Exploitable product(s) or measure(s)	Sector(s) of application[30]	Timetable, commercial or any other use	Patents or other IPR exploitation (licences)	Owner & Other Beneficiary(s) involved
General advancement of knowledge	Phonetic speech recognizer with OOV word detection	NO		Methodology, Software	Information service activities	Public disclosure		BUT
General advancement of knowledge	Large Vocabulary Continuous Speech recognizer for 8kHz	NO		Methodology, Software	Information service activities	Public disclosure		BUT
General advancement of knowledge	Latent topic discovery based on multinomial distribution model (author-post-tag)	NO		Software, Algorithm	Information service activities	Public disclosure		UoKob
General advancement of knowledge	Evolution of latent categories over time	NO		Software, Algorithm	Information service activities	Public disclosure		UoKob

Type of Exploitable Foreground[29]	Description of exploitable foreground	Confidential Click on YES/NO	Foreseen embargo date dd/mm/yyyy	Exploitable product(s) or measure(s)	Sector(s) of application[30]	Timetable, commercial or any other use	Patents or other IPR exploitation (licences)	Owner & Other Beneficiary(s) involved
	using multinomial distribution							
General advancement of knowledge	Low level spam detection method	NO		Method / Software	Information service activities	Public disclosure		CERTH
General advancement of knowledge	Trend Detection framework	NO		Method / Software	Information service activities	Public disclosure		CERTH
General advancement of knowledge	Social Tagging System analysis framework	NO		Method / Software	Information service activities	Public disclosure		CERTH
General advancement of knowledge	Emergency Notification Service	NO		Method / Design / Software	Information service activities	use within organisation		EM-KA
General advancement of knowledge	Relation-based Access Control	NO		Method	Information service activities	use within organisation		EM-KA
General	Community	NO		Design / Software	Information service	use within		EM-KA

Type of Exploitable Foreground[29]	Description of exploitable foreground	Confidential Click on YES/NO	Foreseen embargo date dd/mm/yyyy	Exploitable product(s) or measure(s)	Sector(s) of application[30]	Timetable, commercial or any other use	Patents or other IPR exploitation (licences)	Owner & Other Beneficiary(s) involved
advancement of knowledge	Design Language				activities	organisation		
General advancement of knowledge	Automatic Authorative Role Assignment	NO		Method	Information service activities	use within organisation		EM-KA
General advancement of knowledge	Community Membership Life Cycle Model	NO		Method	Information service activities	use within organisation		EM-KA
General advancement of knowledge	Caching strategies for Access Controls	NO		Software	Information service activities	use within organisation		EM-KA
General advancement of knowledge	Faster modularity clustering algorithm	NO		Algorithm	Information service activities	use within organisation		EM-KA
General advancement of knowledge	Advanced complex Eigenvector centrality	NO		Software / Algorithm	Information service activities	use within organisation		EM-KA

Type of Exploitable Foreground[29]	Description of exploitable foreground	Confidential Click on YES/NO	Foreseen embargo date dd/mm/yyyy	Exploitable product(s) or measure(s)	Sector(s) of application[30]	Timetable, commercial or any other use	Patents or other IPR exploitation (licences)	Owner & Other Beneficiary(s) involved
General advancement of knowledge	Navigation on complex network structures (Community Browser)	NO		Software	Information service activities	use within organisation		EM-KA
General advancement of knowledge	Visualization of community information	NO		Software	Information service activities	use within organisation		EM-KA
General advancement of knowledge	Emergency Alert Service	NO		Software	Information service activities	use within organisation		EM-KA
General advancement of knowledge	Preferred means of communicatio n detection	NO		Method / Software	Information service activities	Public disclosure		CERTH
General advancement of knowledge	Layer and Weight estimation of the relationships between users	NO		Method / Software	Information service activities	Public disclosure		CERTH

Type of Exploitable Foreground[29]	Description of exploitable foreground	Confidential Click on YES/NO	Foreseen embargo date dd/mm/yyyy	Exploitable product(s) or measure(s)	Sector(s) of application[30]	Timetable, commercial or any other use	Patents or other IPR exploitation (licences)	Owner & Other Beneficiary(s) involved
	based on visited content							
General advancement of knowledge	Collective Intelligence methodological approach	NO		Methodology	Information service activities	Public disclosure		CERTH
General advancement of knowledge	Formal representation model of events and multimedia metadata	NO		Method, Design	Information service activities	Public disclosure		UoKob
General advancement of knowledge	Library for using events and multimedia metadata in Java-based applications	NO		Software	Information service activities	Public disclosure		UoKob
General advancement of knowledge	Distributed Group Management	NO		Software, Algorithm	Information service activities	Public disclosure		UoKob

Type of Exploitable Foreground[29]	Description of exploitable foreground	Confidential Click on YES/NO	Foreseen embargo date dd/mm/yyyy	Exploitable product(s) or measure(s)	Sector(s) of application[30]	Timetable, commercial or any other use	Patents or other IPR exploitation (licences)	Owner & Other Beneficiary(s) involved
	based on FOAF profiles							
General advancement of knowledge	Sharing and re-use of social and organisational knowledge for collective intelligence	NO		Methodology	Information service activities	Public disclosure		USFD
General advancement of knowledge	LogMerger tool for searching and merging user logs	NO		Method, Software	Information service activities	Public disclosure		USFD
General advancement of knowledge	SOA architecture solution – based on open-source technologies	NO		Architecture, Software	Information service activities	use within organisation		SMIND
General advancement of	Hardware architecture	NO		Architecture	Information service activities	use within organisation		SMIND

Type of Exploitable Foreground[29]	Description of exploitable foreground	Confidential Click on YES/NO	Foreseen embargo date dd/mm/yyyy	Exploitable product(s) or measure(s)	Sector(s) of application[30]	Timetable, commercial or any other use	Patents or other IPR exploitation (licences)	Owner & Other Beneficiary(s) involved
knowledge	for the WeKnowIt System							
General advancement of knowledge	Guidelines for creation of WKI System architecture compatible services	NO		Method	Information service activities	use within organisation		SMIND
General advancement of knowledge	Web-service interfaces for speech related services	NO		Architecture, Software	Information service activities	use within organisation		BUT
General advancement of knowledge	Development infrastructure (CI server, bugtracker, SCM repository, artifacts repository,	NO		Architecture	Information service activities	use within organisation		SMIND

Type of Exploitable Foreground[29]	Description of exploitable foreground	Confidential Click on YES/NO	Foreseen embargo date dd/mm/yyyy	Exploitable product(s) or measure(s)	Sector(s) of application[30]	Timetable, commercial or any other use	Patents or other IPR exploitation (licences)	Owner & Other Beneficiary(s) involved
	document repository)							
General advancement of knowledge	Java objects (POJOs) storage – part of the WeKnowIt Data Storage	NO		Method, Architecture, Software	Information service activities	use within organisation		SMIND
General advancement of knowledge	Semantic data storage – part of the WeKnowIt Data Storage	NO		Method, Architecture, Software	Information service activities	use within organisation		SMIND
General advancement of knowledge	(Large) Files storage solution – part of the WeKnowIt Data Storage	NO		Method, Architecture, Software	Information service activities	use within organisation		SMIND
General advancement of knowledge	Object translation layer based	NO		Methodology	Information service activities	use within organisation		BUT

Type of Exploitable Foreground[29]	Description of exploitable foreground	Confidential Click on YES/NO	Foreseen embargo date dd/mm/yyyy	Exploitable product(s) or measure(s)	Sector(s) of application[30]	Timetable, commercial or any other use	Patents or other IPR exploitation (licences)	Owner & Other Beneficiary(s) involved
	on @RDF annotations							
Exploitation of results through (social) innovation.	Sheffield floods Emergency Log Sheets	YES		Content	Information service activities	2009	Available through license agreement	SCC
General advancement of knowledge	Web services layer – allows for the communicatio n between the WeKnowIt system and use-case applications.	NO		Architecture, Software	Information service activities	use within organisation		SMIND
General advancement of knowledge	Services composition layer – composes services and performs necessary	NO		Architecture, Software	Information service activities	use within organisation		SMIND

Type of Exploitable Foreground[29]	Description of exploitable foreground	Confidential Click on YES/NO	Foreseen embargo date dd/mm/yyyy	Exploitable product(s) or measure(s)	Sector(s) of application[30]	Timetable, commercial or any other use	Patents or other IPR exploitation (licences)	Owner & Other Beneficiary(s) involved
	data transformation in order to provide the right data to the UI layer of ER use case application.							
General advancement of knowledge	Mobile phone-based interfaces for the speech tagging, tested Android application code	NO		Architecture, Software	Information service activities	use within organisation		BUT
Exploitation of results through (social) innovation	Mobile Guidance Application for travels, enabling recommendati	NO		Software	Information service activities	2011		TID. All consortium also involved

Type of Exploitable Foreground[29]	Description of exploitable foreground	Confidential Click on YES/NO	Foreseen embargo date dd/mm/yyyy	Exploitable product(s) or measure(s)	Sector(s) of application[30]	Timetable, commercial or any other use	Patents or other IPR exploitation (licences)	Owner & Other Beneficiary(s) involved
	ons about Points of Interest, places and events, based on learnt user profiles and considering current geographical location, bookmarks, integrating collective intelligence technologies provided by all project partners							
Exploitation of results through (social) innovation	Enhanced media travel experience –	NO		Software	Information service activities	2011		Yahoo. All consortium also involved

Type of Exploitable Foreground[29]	Description of exploitable foreground	Confidential Click on YES/NO	Foreseen embargo date dd/mm/yyyy	Exploitable product(s) or measure(s)	Sector(s) of application[30]	Timetable, commercial or any other use	Patents or other IPR exploitation (licences)	Owner & Other Beneficiary(s) involved
	both for pre-travel and post-travel scenario.							
Exploitation of results through (social) innovation	Collective Intelligence for Emergency Response			Methodology, Design, Evaluation, Software	Information service activities	2011		USFD. All consortium also involved
Exploitation of results through (social) innovation	WKI Image Recognizer – a mobile application (for Android powered phones) which recognizes pictures taken by user	NO		Software	Information service activities	2011		SMIND, CERTH, YAHOO!

The most obvious way for WeKnowIt partners to exploit the IP is via creation of products and services based on the technology developed in the project. In order for a partner to follow this route, careful assessment of the component parts of the product or subsystem is needed in order to ensure that the partner has exploitation rights for all the necessary elements. In the case of WeKnowIt, all partners contribute to the creation of

specification documents which report on the software and algorithms used in any part of the system. Where a partner makes use of their own background IP, this is made clear, and is covered by the terms of the Consortium Agreement. Therefore, it is possible for partners to exploit the project foreground, in confidence that licences will be made available for use of partner background IP. Where third party IP is brought into the project, this may cause future problems in commercial exploitation if appropriate licences cannot be obtained. Therefore WeKnowIt partners have carried out investigation into commercial exploitation conditions when software is introduced into the project. The specific terms of exploitation of IP are covered by the WeKnowIt Consortium Agreement, which all partners have signed.

Further, licensing of IP is a possible exploitation route for individual component technologies, or groups of technologies, developed in the project, which may be applicable to more than one application or product, which WeKnowIt partners feel can be exploited by other companies without any loss of competitive advantage. Such exploitation is especially advantageous in cases where an unexpected invention or technology breakthrough is achieved, which is peripheral to the main WeKnowIt system development. Licensing is also a beneficial exploitation path for academic partners who do not have their own production and commercialisation facilities, and for industrial partners whose business conditions change during the project's lifetime, such that exploitation no longer fits the organisation's strategic aims. The WeKnowIt Consortium Agreement allows for licensing of IP generated during the project, although specific terms and conditions would be negotiated for each individual licensing opportunity. Licensing is only feasible where the IP has been sufficiently well protected in the project, such that the technology or algorithm is not available to a competitor or potential licensee via another route, and where an obvious work-round to any related WeKnowIt patent does not exist. This implies careful tracking and protection of WeKnowIt IP, for the benefit of the entire consortium, not just the industrial partners.

Finally, the exploitation of IPR generated by academic partners will be mainly on licensing of IPR (mainly software codes and technologies), exploiting in future projects and spin-off company formation to commercialise technologies and IPR generated within the project.