The Software EBox: Integrated Information for Situational Awareness

Naveen Ashish, Jay Lickfett, Sharad Mehrotra and Nalini Venkatasubramanian
University of California, Irvine
Calit2 and Information and Computer Sciences
Irvine, CA 92697, USA
ashish@ics.uci.edu

Abstract— **This paper describes the development of the "EBox", a system that provides integrated access to a wide variety of information sources relevant to providing situational awareness during emergency response situations. The EBox exemplifies a novel information integration approach in which both pre-existing data and real-time information streams are obtained from multiple sources and are rapidly and intelligently ingested and registered to faclilitate efficient and integrated access at a later time. We describe the design and development of the EBox, and the deployment experience in a drill and demonstration exercise. We further describe key technical challenges that remain and our ongoing research efforts in addressing such challenges.**

Keywords-component; ***Situational awareness, data integration, real-time information.***

I. INTRODUCTION

This paper describes the software EBox that is a software solution for information integration in the context of situational awareness systems for emergency response. Today, first-responders, such as firefighters responding to an incident, typically arrive on the disaster scene with typically little useful information regarding the crisis site, buildings and other infrastructure, occupancy or the presence of victims, or even the exact nature of the call [1]. Obtaining such information quickly and in a fashion in which it can be easily used in order to make decisions can be critical to the successful outcome of the emergency response activities, both from the view of neutralizing any hazards, as well as ensuring the safety of both the first responders and any other people in the vicinity. Instead, firefighters today depend heavily upon a variety of locally gathered information and expertise obtained at the time of the incident. For instance, they may ask for details from local people about building occupancy, hazards, exit routes etc. They may observe the locations of triggered alarms or detectors on a centralized alarm panel. Fire departments do frequently have map books which are usually printed books made on an annual or similar basis containing roads, building perimeters, water sources or fire hydrant locations etc. and these may be consulted.

Recognizing the deficiencies with regard to useful information available to decision makers at the time of an incident, a number of companies have begun to contract with individual organizations or with fire departments to conduct surveys of selected building sites [1]. This "pre-planning" data can be a valuable tool for firefighters during response to various potential types of incidents. However, there are many challenges: (a) pre-plan data is static and it is difficult to maintain or republish the data in an updated form, (b) such data is typically limited to infrastructure centric information such as building floor-plans. We believe that there are multiple other forms of information, which if captured and presented appropriately can provide vital situational information. Technology advances have enabled highly instrumented buildings and infrastructures equipped with sensors that can provide information on the structural integrity of the building, environmental conditions within the space (smoke, temperature, humidity) or multimodal situational awareness through video and audio sensors that can help dynamically determine occupancy levels within the building. Leveraging such infrastructure on the fly to provide better awareness poses multiple challenges such as addressing the diversity of the sensors, integration of sensor data with a-priori information and sensor hardware and software interoperability issues. Finally, other locally maintained data may be useful, for instance inventories of hazardous materials location, meeting and activity calendars for rooms in buildings etc., can provide additional situational awareness information.

To overcome the limitations we observe with custom built solutions namely the high cost and restriction to pre-assembled information, as well as provide the capability to exploit additional information such as local information and data from sensors, we propose a system called the software EBox. The EBox is being implemented as a web-service that enables organizations to provide their available data and information sources, which can then be accessed in an integrated fashion during a response situation. We envision that much of this data is provided in advance. Firefighters will connect to this service both prior to departing from the fire station and while at the crisis site, and download the necessary data into their own information systems in order to help them perform their duties. In addition, the service may offer mechanisms for firefighters to actively control elements of the building's infrastructure, e.g., sensors or surveillance cameras, through the EBox system.

[1] The needs are similar for other first-responders, e.g. law enforcement, however for the purposes of our project we constrain the problem to firefighting

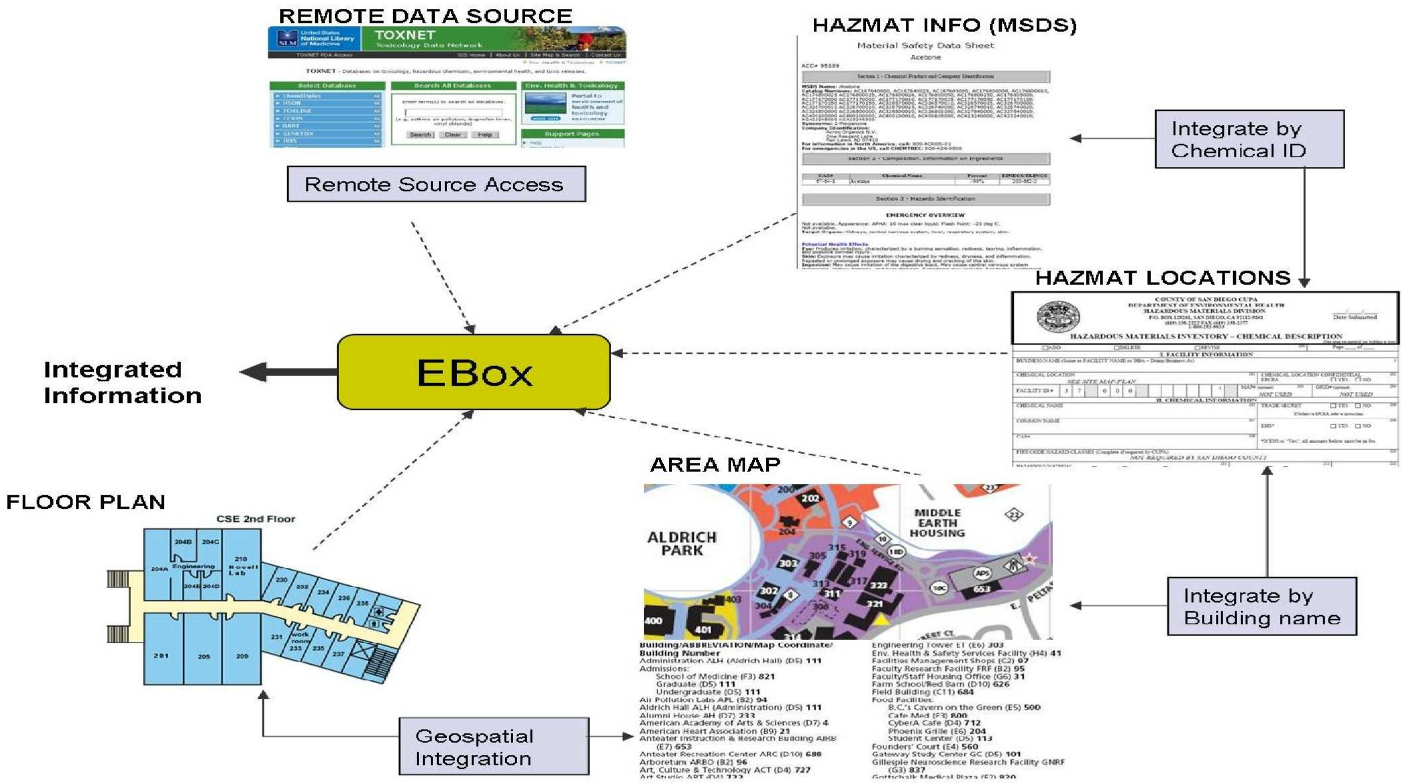

Fig 1 Integration Requirements

The EBox can be viewed as a software and information analog to the traditional concept of a "knox-box" - a small safe located outside a building holding its master keys so that responders can quickly obtain and use them in a response situation.

II. APPLICATION NEED

Consider the following scenario and sequence of events to illustrate the role and value of the EBox. Several calls are received at the 9-1-1 call center reporting a fire in a chemistry lab on the university campus. While fire and hazmat resources are enroute to the scene, the dispatchers check for and activate EBox resources pertaining to the particular location and type of incident. The unlocked EBox resources are available to incident commander and others via networked terminals installed in the fire apparatus and command vehicles.

In this case, the resources available via the EBox could include a) detailed floor plans of the relevant building, b) an up-to-date inventory of hazardous materials obtained live via campus chemical inventory database and cross referenced to the floor plan, c) up-to-date contact phone numbers for the lab managers for the building so that the fire department can reach someone for confirmation of information, d) connection to the building surveillance cameras allowing video feeds to be observed from inside the building, and e) connections into the building alarm system. Even as the firefighters are arriving on the scene and within the first few minutes, the incident commander may already have been able to make several important determinations. The alarm panel indicates the particular lab where the fire is located, and using the phone contact information the incident commander is able to speak with the lab manager. Without this contact information it might have taken much longer to track down the responsible party. The incident commander is able to confirm the presence of a water reactive chemical stored in the lab which is indicated via the EBox hazmat inventory information. This precludes the use of water to extinguish the fire. The lab manager also states that all of the graduate students working in the lab at the time of the fire starting have been accounted for. Depending on the level of instrumentation of a particular building and the sophistication, it would be possible even to review video footage prior to the start of the incident if so allowed by the EBox policies- for example to view exactly how the incident began.

This information allows the incident commander to make several decisions. He decides not to risk entry of the lab area to check for victims because of multiple reports of "all accounted for". Instead he allows the hazmat resources the few extra minutes necessary to don their extra protective suits before entering the hazardous area. Using the detailed floor plans and video cameras he monitors the extent of the fire and can guide firefighters quickly through the relatively safe areas of the building to confirm complete evacuation. In this scenario, use of the EBox has improved the response in several ways. It has sped the ability of the incident commander to reach a relevant point of contact for the area. It has prevented unnecessary exposure of unprotected firefighters to the hazardous environment in the lab fire area, where they otherwise would have looked for victims which were not present. It provides critical information needed to respond to the hazard (i.e., the presence of water reactive chemicals in substantial quantities). Finally it provides a means to utilize observational capabilities of the building infrastructure itself (alarm panel, cameras) to be an additional view into what is actually happening. We note that the EBox is an assistive technology intended to provide the incident commander information useful for decision making and is not intended to automate or remove need for an incident commander or other decision maker to use their own judgement, training, and experience in making such decisions. From a capability perspective, the EBox provides seamless one-stop access to information integrated and over-layed from multiple (possibly independently created) information sources. It provides intelligent search, browsing, and query interfaces so as to be able to search and access information effectively and rapidly in response situations. Fig 1 provides a schematic illustration of the information integration needs in such an application.

III. EBOX DESIGN AND IMPLEMENTATION

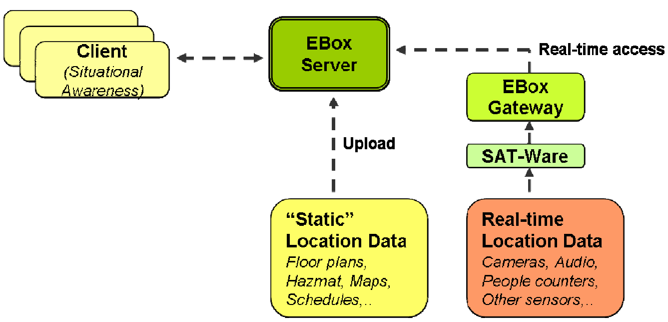

Fig 2 EBox System Architecture

In this section, we describe the design and implementation of the EBox system that we have built. A schematic architecture of the system is provided in Fig 2. The architecture is based upon the "Software-as-a-Service" (SaaS) paradigm where an EBox server can serve particular clients that require information. An EBox "user" i.e., a facility that will use the system first provides available data to the EBox server. This includes data relevant to that facility and the location such as maps of the location and facility, floor plans of buildings in the area, hazardous material location information, and other potentially useful information such as work schedules and shift timing information. All such data is loaded into the EBox server in advance. Further, at the time of an incident and response, we also have the capability of making real-time sensor data available to the EBox server. An example of such a sensor is surveillance cameras in a building at the incident site, which can be integrated in at real-time. The EBox server provides integrated "static" data which is uploaded in advance as well as real-time sensor data to clients i.e., situational awareness systems that will use such data for decision making.

The EBox architecture and implementation strives to achieve two key capabilities for different categories of EBox users, (i) We want to make it easy for a particular facility to provide their available data to the EBox server, (ii) To clients, the EBox must provide effective querying and information retrieval capabilities over the integrated information. Further, the EBox server must facilitate the integration and access of real-time sensor resources.

These capabilities are achieved with many detailed components or modules of the overall system (not illustrated) but which we will briefly describe. These are: **(i) An Ontology Manager:** This module manages the task of creating and maintaining ontologies in EBox applications. It includes some basic "upper" ontologies that are generic across applications. It further provides an interface and GUI tool to a user for creating and extending ontologies specific to applications. The ontology manager also provides interfaces to internal modules that use the ontologies in information organization and retrieval tasks, as we shall elaborate on shortly. **(ii) A Data Ingest Module:** This module handles the task of inserting data into the data repository. Through this module the user provides for storage various kinds of data ranging from maps and floor plans to information about hazardous materials or building occupancy. The user also provides key meta-data for each piece of information that is then used in data organization and for data retrieval. This data ingest module actually uses the domain ontologies to guide a user in providing their data and providing the appropriate meta-data. **(iii) A Data Repository:** This is the repository for all of the pre-assembled data in any EBox application. Storage is provided in a geospatially-enabled relational database as well as a file system. **(iv) The Core EBox Server:** Clients interface with the core EBox server which in turns communicates with other modules in the system to address the client information retrieval requests. **(v) An Ontology Store:** This is the store for all of the pre-assembled as well as newly created or extended ontologies in any application. **(vi) A Search Module:** Provides semantic as well as geospatially and temporally aware search capabilities over the information. We will elaborate on this later. **(vii) An Internet Data Access Module:** There are also many publicly available sources (over the Internet) that provide valuable information. For instance there are online databases of toxicology information or information about chemicals or explosives. The EBox has a module to be able to access and integrate data from such internet sources at real-time. In addition, EBox utilizes a system called **"SAT-Ware"** for interfacing to and integrating real-time sensor data. SAT-Ware [2] is a middleware system for high level access to multi-modal, distributed sensors. While SAT-Ware itself is a general framework developed prior to EBox, we have developed an "EBox-SAT-Ware Gateway" through which the EBox interfaces to SAT-Ware. The EBox creates a representation of each real-time sensor that is available at the incident location. A client request for real-time data from any such sensor (such as say a request to view the video stream from a particular camera) that arrives to the EBox server, is eventually handled through the gateway to SAT-Ware.

In creating any new EBox application or instantiation we go through the following preparation steps using the above modules: (i) *Create Ontologies* The first step is to create and assemble ontologies such as for information categories, key locations in the area of interest, and also (generic) spatial and temporal concepts and relationships. (ii) *Data Ingest* The next step is the ingestion of various available data which is the process of providing data and data source descriptions and meta-data for the information in the various sources. (iii) *Geospatial Anchoring and Integration* Another key step is the geospatial alignment of information from different sources for the purpose of integration. We must ensure that references to locations (such as “Bren Hall” or “Engineering Tower”) are correctly anchored to relevant geography. Further, if multiple imagery datasets or other geographic data of the area are being used (for instance a campus map combined with floor plans for particular buildings) then such geographic data must be registered and projected correctly. (iv) *Integration of Sensor Data* Real-time sensors if available are integrated through the EBox-SAT-Ware gateway. (v) *Integration of Remote Data Sources* We provide real-time query access to any relevant remote (internet) sources through “wrappers” around these information sources.

We also provide some details on the implementation specifics. The Ontology Manager is implemented in Java and provides the Protégé tool to the user for the creation of ontologies. The ontologies are stored in the RDF format. The ontology manager further utilizes Jena as the ontology store and manager. Jena provides for the storage of ontologies in RDF format as well as providing an API for ontology manipulation and querying. The Data Repository is implemented using the MySQL relational database, which includes basic features for storing geographic information. Geographic data is prepared for the system separately, using external GIS tools to geo-register and re-project the data when necessary (such as ArcGIS [16], Manifold [17], GRASS, or GDAL). For the sake of compatibility, all GIS data is required to be in certain specific common projections. Raster imagery is expected to be provided in Universal Transverse Mercator WGS84, and vector data is expected to be in lat/long WGS84. The Data Ingest module is implemented in Java and it interacts with the Data Repository using JDBC. The Central Manager is implemented in Java. The Data Access module uses AJAX and Apache for offering the EBox as a Web service. Finally, the information search functionality is implemented using Lucene.

IV. INSTANTIATION AND VALIDATION

We assembled an EBox instantiation in the context of a larger drill and demonstration exercise. The "incident" triggering the evacuation was a mock hazardous material spill in a hallway of a campus building. During this drill the EBox system was deployed at a simulated incident command post. From this position the incident commander had access to a variety of resources from the EBox. These included detailed floor plans of the two affected buildings, Word documents containing hazmat inventory for nearby labs in one of the buildings, and ability to view video streams from surveillance cameras installed in the two affected buildings. The EBox data was integrated into a separate fire IT system providing an integrated view of the situation to the incident commander including other sensor information and firefighter localization. The drill involved EHS (Environmental Health and Safety) personnel from UCI, observers from the OCFA (Orange County Fire Authority) and City of Ontario fire departments, and the technical team members. We provide details of some of the EBox related information sources that we assembled.

Fig 3 Information Categories Ontology

Ontology Development We developed three ontologies for this particular application namely, (i) An information categories ontology that categorizes the variety of information made available by the EBox, (ii) A locations ontology for the UCI campus capturing significant areas of interest, buildings in these areas etc., and (iii) Ontologies capturing spatial and temporal concepts and relationships. These ontologies were developed using Protégé and a snapshot of the information categories ontologies is provided in Fig 3. The information categories ontology for this application contains about 50 concepts. The ontology further captures the “class-subclass” and “related-to” relationships amongst various concepts.

Information Sources We integrated the following information and information sources, a) A campus map of UCI, b) Information sheets containing information about the presence of hazardous materials and chemicals in various labs in the UCI Engineering buildings, and c) A file folder containing related MSDS (Materials Safety Data Sheets) sheets in PDF format. We further provided access to camera sensors on the Calit2 building entrance and also provided access to TOXNET[1] which is a Web-based database system of toxicology and hazardous chemicals maintained by the NIH (National Institute of Health). For maintaining meta-data about the information sources and the information content we created 9 different tables in the relational database corresponding to different information sources.

Validation As improving situational awareness (SA) is one of the key intended functionalities of the EBox, we conducted an evaluation to assess the improvement in SA achieved with the use of the EBox system. We follow the SAGAT [10] methodology for evaluating SA where the basic idea is to “freeze” a drill or simulation at certain points. Some of the drill actors are then provided questionnaires at each of these points. These questionnaires contain questions related to all 3 SA levels namely perception, comprehension, and projection. Answers to these questions are provided for the both the cases

[1] http://toxnet.nlm.nih.gov

of having access to the SA system, and not. The comparison of answers to the questions in the two cases then provides an assessment of the impact of the SA system.

Level 1 (Perception)
1. What are the locations for which we have a current alarm ?
2. What is the contact information for personnel such as the lab PI ?
3. What are the hazardous materials and chemicals present at or near the site of interest ?
4. What are the significant properties of the chemicals present ?
5. What is the occupancy state of site of interest ?
6. Is there any real-time footage available for this site ?

Level 2 (Comprehension)
1. Do we have knowledge of threats due to hazardous materials and chemicals ?
2. Do we have the most up-to-date information on hazardous materials and chemicals ?
3. Is the site completely clear of any potential victims ?
4. Do we have exact knowledge the fire extent within the building ?

Level 3 (Projection)
1. Can we eliminate the need to initiate a victim search ?
2. Are chemical threats inside significant enough to order fire fighters to first don protective equipment ?

Table 1. SAGAT Evaluation Questionnaire

A sampling of the questionnaire employed at all 3 levels is provided in Table 1 above. In this scenario the use of the EBox significantly improved SA at all 3 levels at various stages. The improved SA was in terms of better knowledge about hazardous chemicals at the site and their properties and more accurate knowledge about building occupancy. This improved SA ultimately resulted in better decisions by the incident commander, namely that of prioritizing donning protective suits by fire-fighters (important given the presence of particular chemicals at the site) over an immediate victim search as multiple information sources confirm that the lab and building have indeed already been evacuated.

As the system evolves and the capabilities are completely implemented, we will conduct further evaluations of the effectiveness of the system from a data integration perspective. Particularly we will assess the ease and savings in time in developing integrated applications with the EBox system as compared to the existing state of the art.

V. INFORMATION INTEGRATION CHALLENGES

While pervasive and ubiquitous computing systems [12,13] have been proposed for fire-fighting awareness, the space of information integration systems in this domain has primarily been investigated in the commercial space. As opposed to such custom built commercial solutions our aim with the EBox is to provide a system and service using which organizations can rapidly provide and assemble available information that can be used in response in an integrated manner. The technical area of data integration has been extensively investigated in the research community for almost two decades now and the corresponding commercial space - namely that of Enterprise Information Integration (or "EII") [7] has also seen much activity in terms of data integration software product offerings. This work has resulted in general purpose data integration engines for integrated multiple databases and other data sources. More recent work has provided solutions in geospatial data integration [3], "loosely" structured information integration [7], peer-to-peer integration systems [5], and scalable "pay-as-you-go" [4] approaches to information integration on a very large scale. The Semantic-Web [11] effort made further contributions in terms of standard languages and representations (such as RDF and OWL) for the representation of ontologies that provide the semantic glue across multiple information sources.

While many of the approaches and techniques developed for particular tasks are applicable to information integration problems in an EBox setting, we find that such an integration application also has some unique characteristics. The first key distinguishing feature is that the *domain* of information is essentially the same across different instantiations of the EBox. In the case this domain is that of fire-fighting response situational awareness. This leads us to investigate whether the uniformity of the domain, across instantiations, is in fact something that can be exploited for application builders and users. Next, there is a significant aspect of geospatial data integration especially at a *deep* level. One may need to integrate or overlay maps of different kinds of the area or overlay maps with other information such as building footprints, floor-plans of various levels, or say information about the presence of hazardous materials at various particular locations. Finally, the user requires multiple capabilities such as being able to browse, perform an exploratory search, or issue formal structured queries to the system for information.

Based on the above requirements for integration, there are the following directions for further research.

Architecture We are developing an information integration architecture that significantly exploits the uniformity of the domain across instantiations. One of the key things the domain uniformity enables is the potential for re-use across instantiations, specifically we are developing a principled approach to facilitate the re-use of a) Domain ontologies, b) Information itself, and c) Access mechanisms to remote information sources (such as wrappers), across different EBox instantiations. Our approach includes mechanisms for a new instantiation builder to browse and search for such existing resources, and further extend or customize them for their particular instantiation. Another advantage that can be derived is to make the integration system *adaptive* in that we can improve the system from instantiation to instantiation. A post execution analysis of instantiations can provide valuable insights into multiple aspects such as the design of ontologies (such as what concepts are commonly used vs those that are esoteric), frequently accessed data and data sources, inconsistencies that may exist perhaps due to data not having been updated, etc. On the data update issue in particular, we are developing techniques for setting up optimal polling and alerting schedules (to information providers) for the consistency of different information pieces. The approach takes into account factors such as the nature of the data and data source, expectation of update frequency, etc., and attempts to achieve the objectives of both data consistency as well as not inundating information providers with update requests. A somewhat orthogonal aspect in such an integration environment is that the information providers are typically willing to be co-operative in facilitating the integration, in that they are likely willing and motivated to devote at least some time and effort towards ensuring the integration. We are thus also developing tools for assisting information providers, as an example annotation tools that utilize (existing) domain ontologies and using which information providers can mark-up and provide meta-data for their particular data and sources.

Integration We will develop techniques for the automated deep level integration of information with geospatial aspects. One key aspect in achieving automated alignment of information across different maps, building footprints or floor maps, and other descriptive information provided with reference to entities such as buildings or other spaces within an area, is the identification of "control points" across the different sources. While techniques have been developed for automatic identification of control points in particular contexts such as that of road maps [3] we need techniques that can address a wider variety of information sources. Further, we need to be able to accurately geo-code very localized references such as particular buildings or areas within an organization site.

Semantic and "Situationally Aware" Search We envision the use of multiple ontologies with linkages across ontologies in any instantiation. While generic 'semantic search' [14] techniques that can exploit such ontological concepts and relationships in searching for information exist, there is a need to effectively handle search terms important from a situational awareness perspective. These include geospatial search terms (for instance "near" or "around") and temporal search terms (such as "most recent") in search queries. While there is work in area such as geospatially aware search technologies [15], we need comprehensive search solutions from a situational awareness perspective.

We described above the motivation for the information integration EBox and presented design and implementation details. We are currently continuing the system development and enhancement, and further working on the selected research areas described above. We thank Paul Amyx of ImageCat Inc., for his guidance on the GIS aspects of this work.

References

[1] [The Tactical Survey Group. http://www.tacticalsurveygroup.com/

[2] Bijit Hore, Hojjat Jaffarpour, Ramesh Jain, Shengyue Ji, Daniel Massaguer, Sharad Mehrotra, Nalini Venkatasubramanian and Utz Westermann. "SATware: Middleware for Sentient Spaces", in "Multimodal Surveillance: Sensors, Algorithms and Systems", 2007.

[3] Ching-Chien Chen, Craig A. Knoblock, Cyrus Shahabi: Automatically and Accurately Conflating Raster Maps with Orthoimagery. GeoInformatica 12(3): 377-410 (2008)

[4] Anish Das Sarma, Xin Dong, Alon Y. Halevy: Bootstrapping pay-as-you-go data integration systems. SIGMOD Conference 2008: 861-874

[5] Alon Y. Halevy, Zachary G. Ives, Jayant Madhavan, Peter Mork, Dan Suciu, Igor Tatarinov: The Piazza Peer Data Management System. IEEE Trans. Knowl. Data Eng. 16(7): 787-798 (2004)

[6] Jayant Madhavan, Philip A. Bernstein, AnHai Doan, Alon Y. Halevy: Corpus-based Schema Matching. ICDE 2005: 57-68

[7] Alon Y. Halevy, Naveen Ashish, Dina Bitton, Michael J. Carey, Denise Draper, Jeff Pollock, Arnon Rosenthal, Vishal Sikka (2005). "Enterprise information integration: successes, challenges and controversies". SIGMOD 2005: 778-787.

[8] Daniela Florescu, Alon Y. Levy, Alberto O. Mendelzon: Database Techniques for the World-Wide Web: A Survey. SIGMOD Record 27(3): 59-74 (1998)

[9] Naveen Ashish. Optimizing Information Mediators by Selectively Materializing Data. PhD Dissertation, University of Southern California, 2000.

[10] Mica Endsley. Direct measurement of situation awareness: Validity and use of SAGAT. In M. R. Endsley & D. J. Garland (Eds.), Situation awareness analysis and measurement. Mahwah, NJ: LEA. (2000)

[11] Tim Berners-Lee, Jim Hendler, and Ora Lasilla. The Semantic Web Scientific American 2001 284(5): 34–43

[12] X. Jiang, N. Y. Chen, J. I. Hong, K. Wang, L. Takayama and J. A. Landay, "Context-aware computing for fire-fighting," in *Pervasive Computing,* 2004,

[13] X. Jiang, J. I. Hong, L. Takayama and J. A. Landay, "Ubiquitous computing for fire-fighters: Field studies and prototypes of large displays for incident command," in *ACM CHI,* 2000

[14] Ramanathan V. Guha, Rob McCool, Eric Miller: Semantic search. WWW 2003: 700-709

[15] C.B. Jones, A.I. Abdelmoty, D. Finch, G. Fu, S. Vaid (2004). "The SPIRIT Spatial Search Engine: Architecture, Ontologies and Spatial Indexing". – 3rd Gi Science 2004 Adelphi, Md, Usa, October 20-23, 2004 Proceedings", pp. 125 - 139.

[16] ArcGIS http://www.esri.com/software/arcgis/

[17] Manifold http://www.manifold.net