Live Geography – Embedded Sensing for Standardised Urban Environmental Monitoring

Bernd Resch [1,2]
Research Scientist
berno
[at] mit.edu
Member, IEEE

Manfred Mittlboeck [1]
Key Researcher
manfred.mittlboeck
[at] researchstudio.at

Fabien Girardin [2,3]
Researcher
fabien.girardin
[at] upf.edu

Rex Britter [2]
Visiting Professor
rb11
[at] eng.cam.ac.uk

Carlo Ratti [2]
Director and
Associate Professor
ratti
[at] mit.edu

[1] Research Studios Austria
studio *i*SPACE
Leopoldskronstrasse 30
5020 Salzburg, Austria

[2] MIT
SENSEable City Lab
77 Massachusetts Avenue
building 10, room 400
Cambridge, MA 02139, USA

[3] Universitat Pompeu Fabra
Department of Information and
Communication Technologies
Passeig de Circumval.lació 8
08003 Barcelona, Spain

***Abstract* – Environmental monitoring faces a variety of complex technical and socio-political challenges, particularly in the urban context. Data sources may be available, but mostly not combinable because of lacking interoperability and deficient coordination due to monolithic and closed data infrastructures. In this work we present the *Live Geography* approach that seeks to tackle these challenges with an open sensing infrastructure for monitoring applications. Our system makes extensive use of open (geospatial) standards throughout the entire process chain – from sensor data integration to analysis, Complex Event Processing (CEP), alerting, and finally visualisation. We discuss the implemented modules as well as the overall created infrastructure as a whole. Finally, we show how the methodology can influence the city and its inhabitants by „making the abstract real“, in other words how pervasive environmental monitoring systems can change urban social interactions, and which issues are related to establishing such systems.**

Keywords – Urban environmental monitoring; Standardised infrastructure; Real-time GIS data analysis; Situational awareness; Embedded sensor device.

I. INTRODUCTION

Environmental monitoring is a critical process in cities to ensure public safety including the state of the national infrastructure, to set up continuous information services and to provide input for spatial decision support systems. However, setting up an overarching monitoring system is not trivial. Currently, different authorities with heterogeneous interests each implement their own monolithic infrastructures to achieve very specific goals, as stated in our previous work [1]. For instance, regional governments measure water levels for flood water prediction, while local governments monitor air quality to dynamically adapt traffic conditions, and energy providers assess water flow in order to estimate energy potentials.

The fact that these systems tend to be deployed in an isolated and uncoordinated way means that the automatic assembly and analysis of these diverse data streams is impossible. However, making use of all available data sources is a prerequisite for holistic and successful environmental monitoring for broad decision support in an urban context. This applies to emergency situations as well as to continuously monitoring urban parameters.

One way to overcome this issue is the extensive use of open standards and Geographic Information System (GIS) web services for structuring and managing these heterogeneous data. Here, the main challenge is the distributed processing of vast amounts of sensor data in real-time, as the widespread availability of sensor data with high spatial and temporal resolution will increase dramatically with rapidly decreasing prices [2], particularly if costs are driven down by mass utilisation.

From a political and legal standpoint, national and international legislative bodies are called upon to foster the introduction of open standards in public institutions. Strong early efforts in this direction have been made by the European Union (EU) through targeted directives (s. chapter IV). These regulations support the development of ubiquitous and generically applicable real-time data integration mechanisms. Shifting development away from proprietary single-purpose implementations towards interoperable analysis systems will not only enable live assessment of the urban environment, but also lead to a new perception of the city by its inhabitants. Consequently, this may in turn foster the creation of innovative applications that treat the city as an interactive sensing platform, such as *WikiCity* [3], involving the people themselves into re-shaping the urban context.

This paper begins with a review of related work in several research areas. Then, challenges of environmental monitoring with particular respect to the urban context are elucidated, before we summarise the current legal frameworks for environmental data management. Thereafter, our *Live Geography* approach is presented, which aims to integrate live sensor measurements with archived data sources on the server side in a highly flexible and interoperable infrastructure. Finally, we present our thoughts on how environmental sensing and Geographic Information (GI) processing can affect the city and its inhabitants. The ultimate goal of this paper is to present our approach's potential impact on urban policy and decision-making, and to point out its portability to other application domains.

II. RELATED WORK

The Live Geography approach is manifold in terms of both concepts and employed technologies. As such, there are several research initiatives that form part of the overall methodology. These are described below.

The first domain is **sensor network development for environmental monitoring**. The Oklahoma City Micronet [4] is a network of 40 automated environmental monitoring stations across the Oklahoma City metropolitan area. The network consists of 4 Oklahoma Mesonet stations and 36 sites mounted on traffic signals. At each traffic signal site, atmospheric conditions are measured and transmitted every minute to a central facility. The Oklahoma Climatological Survey receives the observations, verifies the quality of the data and provides the data to Oklahoma City Micronet partners and customers. One major shortcoming of the system is that it is a much specialised implementation not using open standards or aiming at portability. The same applies to CORIE [5], which is a pilot environmental observation and forecasting system (EOFS) for the Columbia River. It integrates a real-time sensor network, a data management system and advanced numerical models.

Secondly, there are a number of approaches to **leveraging sensor information in GIS applications**. [6] presents the SenseWeb project, which aims to establish a Wikipedia-like sensor platform. The project seeks to allow users to include their own sensors in the system and thus leverage the „community effect“, building a dense network of sensors by aggregating existing and newly deployed sensors within the SenseWeb application. Although the authors discuss data transformation issues, data fusion, and simple GIS analysis, the system architecture is not based on open (geospatial) standards, only standard web services. The web portal implementation, called SensorMap, uses the Sensor Description Markup Language (SDML), an application-specific dialect of the Open Geospatial Consortium (OGC) SensorML standard.

In [7], the author presents a sensing infrastructure that attempts to combine sensor systems and GIS-based visualisation technologies. The sensing devices, which measure rock temperature at ten minute intervals, focuses on optimising resource usage, including data aggregation, power consumption, and communication within the sensor network. In its current implementation, the infrastructure does not account for geospatial standards in sensor observations. The visualisation component uses a number of open standards (OGC Web Map Service [WMS], Web Feature Service [WFS]) and open-source services (UMN Map Server, Mapbender).

Another sensing infrastructure is described in [8]. The CitySense project uses an urban sensor network to measure environmental parameters and is thus the data source for further data analysis. The project focuses on the development of a city-wide sensing system using an optimised network infrastructure. An important parallel with the work presented in this paper is that CitySense also considers the requirements of sensor network setup in an urban environment.

A GIS mashup for environmental data visualisation is presented in the nowCOAST application [9]. Data from several public providers are integrated in a web-based graphical user interface. nowCOAST visualises several types of raw environmental parameters and also offers a 24-hour sea surface temperature interpolation plot.

The most striking shortcoming of the approaches described above and other related efforts is that their system architectures are at best partly based on open (geospatial) standards.

The third related research area is **real-time data integration for GIS** analysis systems. Most current approaches use web services based on the classic request/response model. Although partly using open GIS standards, they are often unsuitable for the real-time integration of large volumes of data. [10] establishes a real-time spatial data infrastructure (SDI), which performs several application-specific steps (coordinate transformation, spatial data generalisation, query processing or map rendering and adaptation), but accounts neither for event-based push mechanisms nor for the integration of sensor data.

Other approaches for real-time data integration rely on the costly step of creating a temporal database. Oracle's system, presented in [11], is essentially a middleware between (web) services and a continuously updated database layer. Like Sybase's method [12], the Oracle approach detects database events in order to trigger analytical actions accordingly. In [13], a more dynamic method of data integration and fusion is presented using on-the-fly object matching and metadata repositories to create a flexible data integration environment.

The fourth comprised research field is the development of an **open data integration system architecture** in a non-application-specific infrastructure. Recent research efforts focus on general concepts in systems architecture development and data integration, but there are mostly no concrete conclusions as to how to establish such an infrastructure. A more technical approach for ad-hoc sensor networks is described in [14], where the authors discuss application-motivated challenges to combining heterogeneous sensor measurements through highly flexible middleware components. The method is strongly application-motivated and thus very well-thought-out as far as specific implementation details are concerned.

III. Challenges of Urban Environmental Monitoring and Sensing

The urban context poses many challenges to environmental monitoring: not only are there significant technical and technological issues, but also social and political ones as well.

The key technological challenge is the integration of different data sources owned by governmental institutions, public bodies, energy providers and private sensor network operators. This problem can be tackled with self-contained and well-conceived data encapsulation standards – independent of specific applications – and enforced by legal entities, as discussed in chapter IV. However, the adaptation of existing sensors to new standards is costly for data owners and network operators in the short term, and so increased awareness of the benefits of open standards is required.

From a technical viewpoint, unresolved research challenges for ubiquitous urban monitoring infrastructures are manifold and include: finding a uniform representation method for measurement values, optimising data routing algorithms in multi-hop networks, and developing optimal data visualisation and presentation methods. The last is an essential aspect of decision support systems, as different user groups might need different views of the underlying information. For example, in emergency local authorities might want a socio-economic picture of the affected areas, while first-response forces are interested in topography and people's current locations, and the public might want general information about the predicted development of a disaster.

From a more contextual standpoint, an important peculiarity of the urban context is that there are large variations within continuous physical phenomena over small spatial and temporal scales. For instance, due to topographical, physical or optical irregularities, pollutant concentration can differ considerably, even on opposite sides of the street. This variability tends to make individual point measurements less likely to be representative of the system as a whole. The consequence of this dilemma is an evolving argument for environmental regulations based on comprehensive monitoring data rather than mathematical modelling, and this demand is likely to grow. Consequently, the deployment of many sensors allows for more representative results together with an understanding of temporal and spatial variability.

One way to overcome this issue is to „sense people" and their immediate surroundings using everyday devices such as mobile phones or cameras. These can replace – or at least complement – the extensive deployment of specialised city-wide sensor networks. The basic trade-off of this people-centric approach is between cost efficiency and real-time fidelity. We believe that the idea of using existing devices to sense the city is crucial, but that it requires more research on sensing accuracy, data accessibility and privacy, location precision, and interoperability in terms of data and exchange formats. Furthermore, measurements are only available in a quasi-continuous distribution due to the high spatial and temporal variability of ad-hoc data collection. Addressing this issue will require complex distribution models and efficient resource discovery mechanisms in order to ensure adaptability to rapidly changing conditions.

Another central issue in deploying sensor networks in the city is the impact of fine-grained urban monitoring, as terms like „air quality" or „pollutant dispersion" are only a surrogate for a much wider and more direct influence on people, such as life expectation, respiratory diseases or quality of life. This raises the demand of finding the right level of information provision. More accurate, finer-grained or more complete information might in many cases not necessarily be worthwhile having, as this could allow for drawing conclusions on a very small scale, in extreme cases even on the individual. This again could entail a dramatic impact in a very wide range of areas like health care, the insurance sector, housing markets or urban planning and management.

Finally, some more unpredictable challenges posed by the dynamic and volatile physical environment in the city are radical weather conditions, malfunctioning hardware, connectivity, or even theft and vandalism.

IV. Policy-Frameworks for the Integration of Real-time Sensor Information

As mentioned above, we have seen an explosion of spatial data collection and availability in digital form in the past several years. There are various national and international efforts to establish spatial data infrastructures (SDI) for promoting and sharing geospatial information throughout governments, public and private organisations and the academic community. It is a substantial challenge solving the political, technological and semantic issues for sharing geographic information to support decision making in an increasingly environment-oriented world. In 2007, the United Nations Geographic Information Working Group published a report subsuming recent regional national and international technologies, policies, criteria, standards and people necessary to organise and share geographic information. These include real-time location aware sensor measurements to develop a United Nations Spatial Data Infrastructure (UNSDI) and encourage interoperability across jurisdictions and between UN member states. As described, these SDIs should help stimulate the sharing and re-use of expensive geographic information in several ways:

- The *Global Spatial Data Infrastructure Association* is one of the first organisations to promote international cooperation in developing and establishing local, national and international SDIs through interaction between organisations and technologies supported by the U.S. Geological Survey (USGS).
- On a supra-national level, the *INfrastructure for SPatial Information in Europe* (INSPIRE) aims to enable the discovery and usage of data for analysing and solving environmental problems by overcoming key barriers such as inconsistency in data collection, a lack of documentation, and incompatibility between legal and geographic information systems.

- *Global Monitoring for Environment and Security* (GMES) is another European Initiative for the implementation of information services dealing with environmental and security issues using earth information and in-situ data for the short, mid and long-term monitoring of environmental changes. This is Europe's main contribution to the Group of Earth Observations (GEO) for monitoring and management of planet earth.
- *Global Earth Observation System of Systems* (GEOSS) seeks to establish an overarching system on top of national and supra-national infrastructures to provide comprehensive and coordinated earth observation for transforming these data into vital information for society.

The common goal of all these initiatives – and the numerous national SDI approaches – is the integration and sharing of environmental information comprising remote sensing data, geographic information and (real-time) measurement data sets. With the European Shared Environmental Information System (SEIS), a new concept has been introduced to collect, analyse and distribute these information sets in a loosely coupled, „federal" structure focused on defining interfaces. A particular focus is dedicated to near real-time datasets like sensor measurements. This new flexibility should foster the integration of sensor measurements into existing SDIs.

V. LIVE GEOGRAPHY APPROACH

With the above mentioned challenges to urban monitoring and standardisation in mind, we have created the *Live Geography* approach, which aims to combine live measurement data with historic data sources in an open standards-based infrastructure using server-side processing mechanisms.

The system architecture is composed of layers of loosely-coupled and service-oriented building blocks, as described in the following sections. In this way, data integration can be decoupled from the analysis and visualisation components, allowing for flexible and dynamic service chaining. In order to fulfil real-time data needs and alerting requirements, the concept also incorporates an event-based push mechanism (sub-section B). As noted above, one of the major challenges is the integration of location-enabled real-time measurement data into GIS service environments to perform distributed analysis tasks. There are three main requirements for quality-aware GIS analysis: *accuracy*, *completeness* and *topicality* of the input data (layers). As recent developments often do not account for time stamp parameters, it is necessary to identify effective ways to combine space and terrestrial real-time observation data with SDI information layers.

Fig. 1 illustrates the basic service infrastructure of the Live Geography concept. The general workflow within the infrastructure can be followed from left to right. First, heterogeneous data sources, such as sensor data, external data (provided via standardised interfaces such as OGC WFS or WCS [Web Coverage Service]) or archived data are integrated on the server side. This integration can happen via both classical request/response models and push services that send out alerts, e.g. if a certain threshold is exceeded. The flexible sensor fusion mechanism supports as well mobile as static sensors. Next, the different kinds of data are combined by a data integration server. This step requires real-time processing capabilities, such as Event Stream Processing (ESP) and Complex Event Processing (CEP). The harmonised data are then fed to pre-defined GIS process models to generate user-specific output.

This approach shifts resource-consuming geo-processing operations away from the client by executing complex, asynchronous analysis tasks on the server side, and then simply providing the client with a tailored result. The output could be an XML structure, a numerical value, or a contextual map tailored to the user's specific needs. The crucial benefit of this approach is that GIS applications – that previously offered GIS functionality only through resource-consuming desktop clients – could be replaced by lightweight web-based analysis tools. This enables the results of GIS analysis to be delivered to a wide variety of internet-connected devices, including personal computers, handhelds, and smart phones, or even other online analytical processes. This also allows for real-time situational awareness in spatial decision support systems. In other words, the system is suitable for using GIS-compliant data sets to assess urban environmental conditions and predict, within limits, their development in real-time. [15]

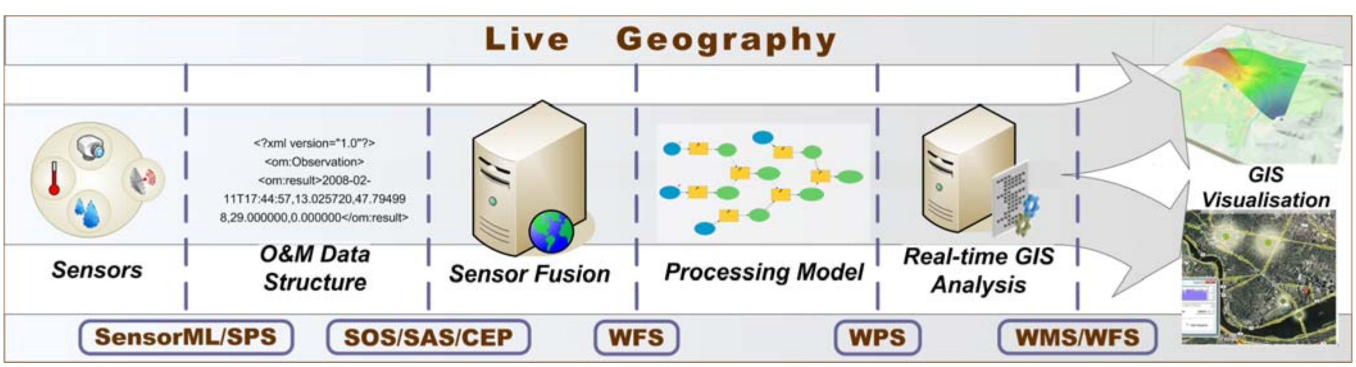

Figure 1. Live Geography Infrastructure.

A. *Usage of Open Standards*

The components of this process chain are separated by several interfaces, which are defined using open standards. The first central group of standards is subsumed under the term Sensor Web Enablement (SWE), an initiative by the OGC that aims to make sensors discoverable, query-able, and controllable over the Internet [16]. Currently, the SWE family consists of seven standards, which encompass the entire process chain from making sensors discoverable in a registry, to measuring physical phenomena, and sending out alerts. [17]

- *Sensor Model Language (SensorML)* – This standard provides an XML schema for defining the geometric, dynamic and observational characteristics of a sensor. Thus, SensorML assists in the discovery of different types of sensors, and supports the processing and analysis of the retrieved data, as well as the geo-location and tasking of sensors.
- *Observations & Measurements (O&M)* – O&M provides a description of sensor observations in the form of general models and XML encodings. This framework labels several terms for the measurements themselves as well as for the relationship between them. Measurement results are expressed as quantities, categories, temporal or geometrical values as well as arrays or composites of these.
- *Transducer Model Language (TML)* – Generally speaking, TML can be understood as O&M's pendant or streaming data by providing a method and message format describing how to interpret raw transducer data.
- *Sensor Observation Service (SOS)* – SOS provides a standardised web service interface allowing access to sensor observations and platform descriptions.
- *Sensor Planning Service (SPS)* – SPS offers an interface for planning an observation query. In effect, the service performs a feasibility check during the set up of a request for data from several sensors.
- *Sensor Alert Service (SAS)* – SAS can be seen as an event-processing engine whose purpose is to identify pre-defined events such as the particularities of sensor measurements, and then generate and send alerts in a standardised protocol format.
- *Web Notification Service (WNS)* – The Web Notification Service is responsible for delivering generated alerts to end-users by E-mail, over HTTP, or via SMS. Moreover, the standard provides an open interface for services, through which a client may exchange asynchronous messages with one or more other services.

Furthermore, *Sensor Web Registries* play an important role in sensor network infrastructures. However, they are not decidedly part of SWE yet, as the legacy OGC Catalogue Service for Web (CSW) is used. The registry serves to maintain metadata about sensors and their observations. In short, it contains information including sensor location, which phenomena they measure, and whether they are static or mobile. Currently, the OGC is pursuing a harmonisation approach to integrate the existing CSW into SWE by building profiles in ebRIM/ebXML (e-business Registry Information Model).

An important ongoing effort in SWE development is the establishment of a central *SWE Common* specification. Its goal is to optimise redundancy and maximise reusability by grouping common elements for several standards under one central specification.

The functional connections between the described standards are illustrated in Fig. 2.

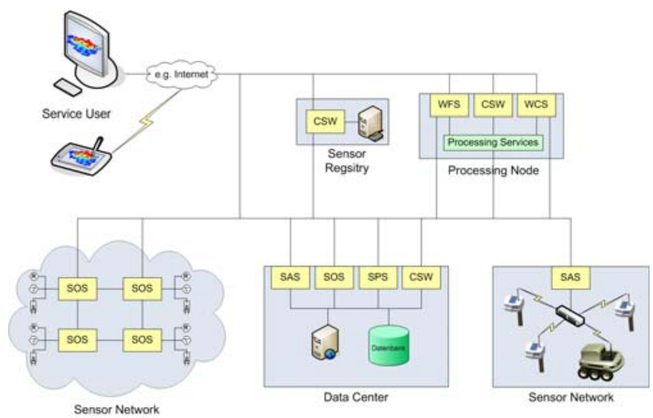

Figure 2. Functional Connections between the SWE Standards.

Besides these sensor-related standards, other OGC standards are used for data analysis and provisioning. The Web Processing Service, as described in [18], provides an interface to access a processing service offering a number of pre-defined analytical operations – these can be algorithms, simple calculations, or more complex models, which operate on geospatial data. Both vector and raster data can be processed. The output of the processes can be either a pre-defined data structure such as Geographic Markup Language (GML), geoRSS, Keyhole Markup Language (KML), Scalable Vector Graphics (SVG), or a web-accessible resource like a JPEG or PNG picture.

Standardised raw data access is granted by the use of OGC WFS, WMS and WCS standards. These well-known standards provide access to data in various formats such as vectors (points, lines and polygons), raster images, and coverages (surface-like structures).

More about the described sensor related and data provision standards can be found on the OGC web site[1].

B. *Location-aware Complex Event Processing*

Apart from standardised data transmission and provision, a special focus in the Live Geography approach is the extension of Complex Event Processing (CEP) functionality by spatial parameters. In a geographic context, CEP can for instance serve for detecting threshold exceedances, for geo-

[1] http://www.opengeospatial.org

fencing implementations, for investigating spatial clusters, or for ensuring data quality.

Generally speaking, CEP is a technology that extracts knowledge from distributed systems and transforms it into contextual knowledge. Since the information content of this contextual knowledge is higher than in usual information, the business decisions that are derived from it can be more accurate. The CEP system itself can be linked into an Event Driven Architecture (EDA) environment or just be built on top of it. While the exact architecture depends on the application-specific needs and requirements, a general architecture including five steps can be applied to every CEP system. Fig. 3 shows this architecture in Job Description Language (JDL) [19].

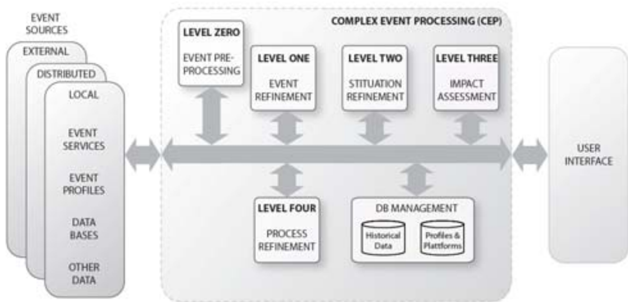

Figure 3. General CEP Architecture with Levels of Data Processing. [19]

The system is divided into several components (or levels) that represent processing steps. Since individual systems may have different requirements the implementation depth will vary from system to system. The following list notes the basic requirements for each level, according to [19]:

- *Level 0: Pre-processing* – Before the actual processing takes part the data is normalised, validated and eventually pre-filtered. Additionally, it may be important to apply feature extraction to get rid of data that is not needed. Although this part is important for the CEP workflow, it is not a particularity of CEP in general. Thus, it is marked as level 0.
- *Level 1: Event Refinement* – This component's task is to track and trace an event in the system. After an event has been tracked down, its characteristics (e.g. data, behaviour and relationships) are translated into event-attributes. The tracing part deals with state estimation that tries to predict upcoming events in the system.
- *Level 2: Situation Refinement* – The heart of each CEP system is the step of situation refinement, where the actual analysis of simple and complex events takes place. This analysis includes mathematical algorithms, which comprise not only computing boundaries for certain values, but also matching them against patterns and historical data. The results are high level (contextual) interpretations which can be acquired in real-time.
- *Level 3: Impact Assessment* – After analysing and refining the situation, it is important to generate decisions that may have consequences. Impact assessment deals with the simulation of outcomes. It therefore deals with various scenarios and simulates them by accounting for cost factors and resources. Results are weighted decision reports that include priorities and proposals for corresponding scenarios.
- *Level 4: Process Refinement* – Finally, the last step covers the interaction between the CEP system and business processes. It provides a feedback loop to control and refine business processes. While this could include integration and automatic controlling of processes, it may also just be the creation of alerting messages or business reports.

Originally, CEP has been developed and traditionally been implemented in the financial and economic sectors to predict market developments and exchange rate trends. In these areas, CEP patterns emerge from relationships between the factors *time*, *cause* (dependency between events) and *aggregation* (significance of an event's activity towards other events). In a location-aware CEP, these aspects get extended by additional parameters indicating *location* information of the event. Spatial parameters are combined on par with other relational indicators.

As geo-referenced data is therefore subject for further processing, one important aspect is to control its data quality. Quality criteria comprise lineage, logical consistency, completeness or temporal quality. From a practical viewpoint, this means that the location may be used to define quality indicators, which can be integrated into CEP pattern rules.

C. Implementation

The implementation of the Live Geography approach comprises tailor-made sensing devices, a real-time data integration mechanism, a use case specific interpolation model, an automatic server-based analysis component, and a complex event processing and alerting component. All these stand-alone modules are chained together using open standards as described below.

For the **measurement device**, we designed a special sensing pod for pervasive GIS applications using ubiquitous embedded sensing technologies. The system has been conceived in such a modular way that the base platform can be used for a variety of sensor web applications such as environmental monitoring, biometric parameter surveillance, critical infrastructure protection or energy network observation by simply changing the interfaced sensors.

The sensor pod itself consists of a standard embedded device, a Gumstix Verdex XM4 platform including an ARM7-based 400MHz processor with 64MB RAM and 16MB flash memory. It runs a customised version of the „Open Embedded“ Linux distribution (kernel version 2.6.21) with an overall footprint of <8MB. Additionally to this basic operating system, we attached a GPS module (U-BLOX NEO 4S and LEA-4P) for positioning and several different sensors (e.g. LM92 temperature, NONIN 8000SM oxygen saturation and pulse, or SSM1 radiation sensors).

The software infrastructure comprises an embedded secure web server (Nostromo nhttpd), an SQLite database and several daemons, which convert sensor readings before they are served to the web. The database serves for short-term storage of historic measurements to allow for different error detection procedures and plausibility checks, as well as for non-sophisticated trend analysis.

For standardised data retrieval, we have created a service implementing the OGC Sensor Observation Service (SOS), SensorML and Observations and Measurements (O&M) standards in an application-specific way. Like this, measurement data are served via HTTP over Universal Mobile Telecommunications System (UMTS) in the standardised XML-based O&M format over the SOS service interface. SensorML serves for describing the whole sensor platform as well as to control the sensor via the OGC Sensor Planning Service (SPS), for instance to dynamically adjust measurement cycles.

The overall service stack on the embedded sensing device is illustrated in Fig. 4.

For the **real-time data integration component**, we have developed a data store extension to the open-source product *GeoServer*[2] 1.6. This plug-in enables the direct integration of OGC SOS responses (in O&M format) into GeoServer and their conversion to OGC-conformal service messages on-the-fly. The main advantage of this approach is that sensor data are made available through a variety of established geo-standardised interfaces (OGC WFS, WMS, WCS), which offer a variety of output formats such as OGC Geographic Markup Language (GML), OGC Keyhole Markup Language (KML), geoRSS, geoJSON, Scalable Vector Graphics (SVG), JPEG or PDF.

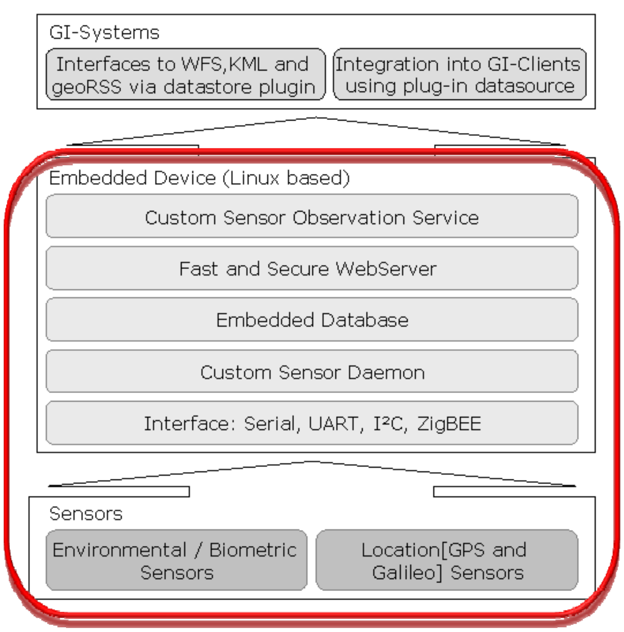

Figure 4. Software Infrastructure on the Embedded Sensing Device.

[2] http://www.geoserver.org

During the transformation procedure from O&M input to WFS output, certain input parameters (coordinate reference system, unit conversions, data structures etc.) are interpreted. In practice, this means that the O&M XML structure is converted into well-established standardised data formats as mentioned above.

The innovation in comparison to previous data integration approaches is that the conversion of the data structure from SOS responses to various WFS, WMS and WCS output formats is performed on-the-fly. Conventional methods typically use the laborious interim step of storing data in a temporary database. This approach has two distinct disadvantages. At first, it adds another component to the overall workflow, which likely causes severe performance losings; secondly, it creates a single central point of failure making the system vulnerable in case of technical malfunctions.

On the contrary, our approach allows for the establishment of a distributed sensor service architecture, and thus enables data provision of real-time measurements for heterogeneous application domains and requirement profiles. The direct conversion of data structures allows for the easy integration of sensor data into GIS applications and therefore enables fast and ubiquitous data visualisation and analysis.

In its current implementation, **geographical analysis** is performed by ESRI's ArcGIS software suite since reliable open processing services are not yet available. We created a *Live Sensor Extension* that allows for the direct integration of standardised sensor data into ArcGIS. The *Live View* component enables ad-hoc GIS processing and visualisation in a navigable map using ESRI's Dynamic Display technology. Fig. 5 shows an interpolated temperature surface in ArcScene using the Inverse Distance Weighting (IDW) algorithm. In addition to ArcGIS, we have also used a number of open/free clients such as Open Layers, uDig, Google Maps, Google Earth or Microsoft Virtual Earth to visualise sensor information in 2D and 3D.

A second processing module implementing the Live Geography framework is an ArcGIS Tracking Analyst based spatio-temporal analysis component. For our study, we used CO_2 data captured by the CitySense [8] network in Cambridge, MA US.

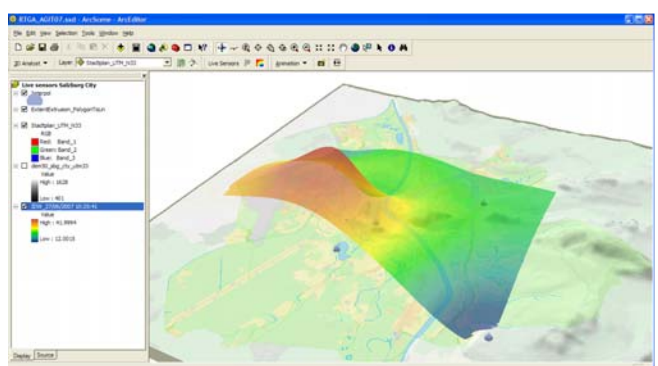

Figure 5. 3D Interpolation Using the Inverse Distance Weighting (IDW) Algorithm.

Fig. 6 shows the interface, which illustrates a time series of measurement data over a period of time. The lower left part of the figure shows the temporal gradient of the measurement values. Running the time series then changes symbologies in the map on the right side accordingly in a dynamic manner. Preliminary findings show that CO_2 is characterised by very high temporal and spatial fluctuations, which are induced by a variety of factors including temperature variability, time during the day, traffic emergence or „plant respiration". In further analysis, this would allow for instance correlating temporal measurement data fluctuation to traffic density, weather conditions or day-time related differences in a very flexible way. Together with the Public Health Department of the City of Cambridge, we are currently carrying out more detailed investigations on these aspects.

A particularly innovative part of the implementation is the web-based GI processing component. We established two data analysis models for Inverse Distance Weighting (IDW) and Kriging operations. Together with the web interface source code itself, we then integrated these models into a single toolbox, which can be published as a web service on ArcGIS server. This allows for data analysis by just selecting base data and the according processing method in a two-click procedure. Two distinct advantages of this mechanism versus current desktop GIS solutions are that geographic analysis can be done without profound expert knowledge, and that processing requirements are shifted away from desktop computers to the server-side. These benefits will likely induce a paradigm shift in GI data processing in the next years and foster a broader spectrum of application areas for spatial analysis operations.

Furthermore, we implemented a **CEP and alerting mechanism** based on XMPP (Extensible Messaging and Presence Protocol), conformant to the OGC Sensor Alert Service (SAS) specification.

In our case, CEP is used for detecting patterns in measurement data and for creating complex events accordingly. These can be related to time (temporal validity), space (e.g. geo-fencing with geographic „intersect", „overlap", or „join" operations), or measurement parameters (e.g. threshold exceedances). Event recognition and processing happens in two different stages of the workflow. Firstly, at sensor level CEP is used to detect errors in measurement values by applying different statistical operations such as standard deviations, spatial and temporal averaging, or outlier detection. Secondly, after the data harmonisation process CEP serves for spatio-temporal pattern recognition, anomaly detection, and alert generation in case of threshold transgression.

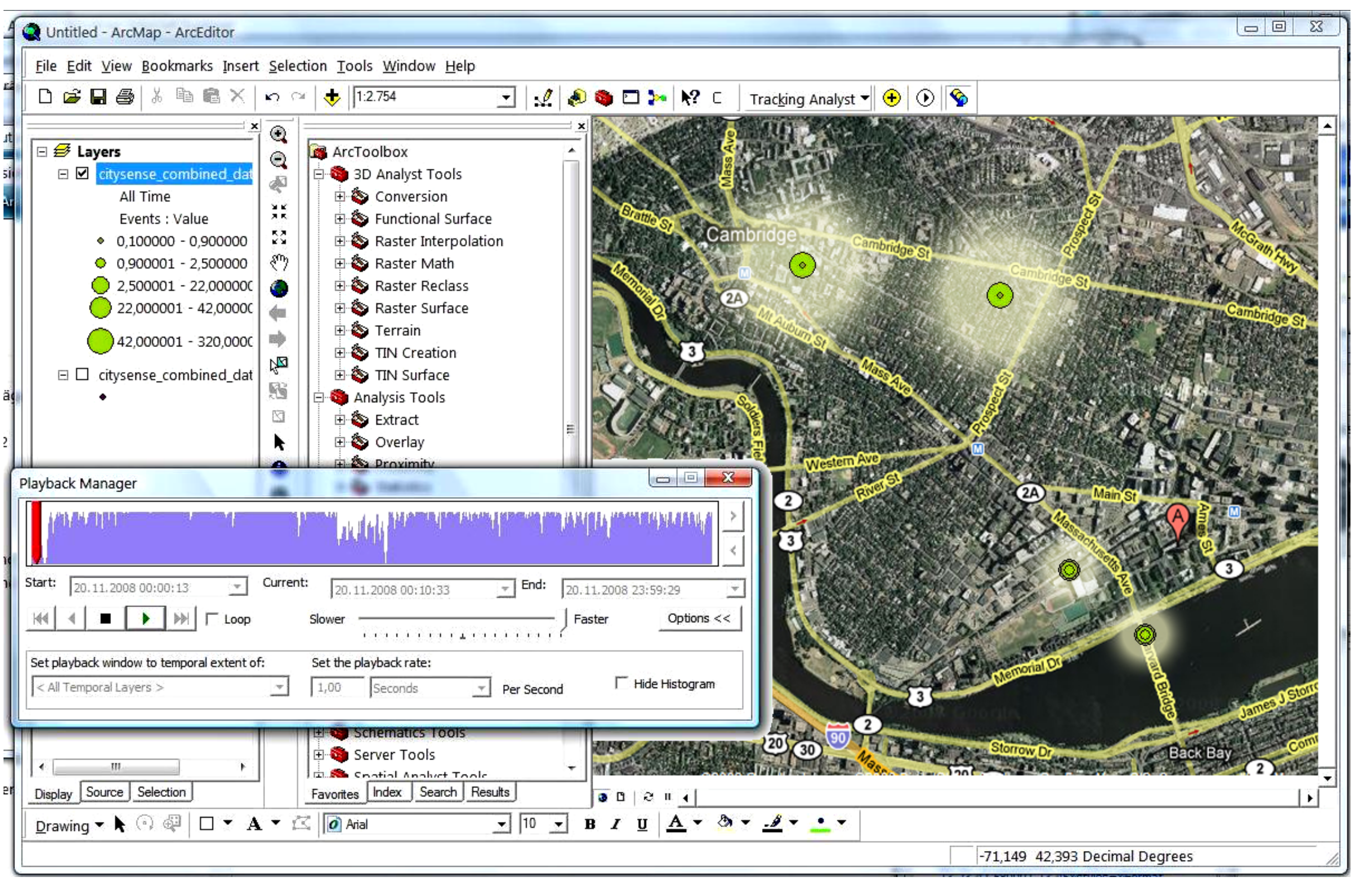

Figure 6. Time Series Analysis in Tracking Analyst.

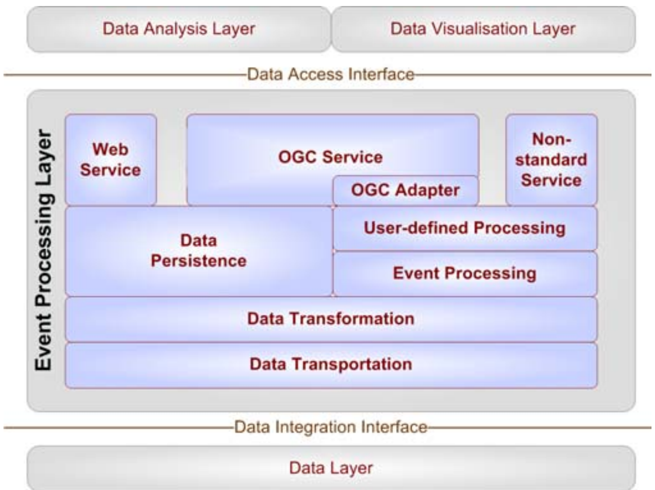

Figure 7. Internal Service Stack of the CEP Component.

In the actual implementation, we used the Esper CEP engine in its version 3.0.0 because of its open availability, Event Query Language (EQL) based event description, and its simple integration by just including a single Java library. For sending events created by the CEP engine, we realised a push-based OGC SAS compliant alerting service. SAS is an asynchronous service connecting a sensor in a network to an observation client. In order to receive alerts, a client subscribed to the SAS. If the defined rules apply, a pre-defined alert is sent to the client via XMPP. It shall be stated that the whole communication between the embedded XMPP server (jabberd2) and the client is XML-based for simplifying M2M messaging.

Fig. 7 shows the internal sub-parts of the CEP-based event processing component, which is built up in a modular structure. Generally speaking, the event processing component connects the data layer (i.e. sensor measurements), and the data analysis and data visualisation components. In other words, it prepares raw data in order to be process-able in the analysis and the visualisation layers. The data transportation component is responsible for connecting the data integration layer to a wide variety of data sources, which can comprise sensor data, real-time RSS feeds, ftp services, web services, databases etc. Hence, it serves as an entry point into the system. Its main responsibility is to receive various kinds of data structures and pass them on as is to the data transportation layer adding metadata do the actual payload, such as the data source or its format.

The event processing component handles the objects coming from the transformation module according to specified user query statements. This module basically handles multiple streams and identifies and selects meaningful events. Currently, the Esper event processing engine is used to detect events and push data to the user processing component. We extended the Esper engine by spatial parameters as described above. Following the event processing step, a further user-defined processing method is applied. This component implements a set of filtering or selection rules, which are specified by the user according to a very specific need.

The data persistence component receives a set of data from the data transformation module or from the processing modules (i.e. a „filtered“ dataset). From these data, it creates a physical data structure, which can either be temporary (for rapidly changing real-time data) or permanent (for time-insensitive data as created by stationary measurement devices). Consequently, as well live data sources as static ones can be handled by the data persistence component.

The non-standard service is one of three service interfaces, which connect the data integration layer to the data analysis layer. It provides data via a custom (i.e. non-standard) interface. This is necessary as existing standardised OGC services do not automatically support push mechanisms. It shall be mentioned that standardised data structures can also be served via the non-standard interface, just the service interface itself is not wholly standards-conformal. The service is also responsible for converting objects, which are created by the data transformation component, to a common, pre-defined output format. The output can either be GML, KML or geoRSS for spatial data, or RSS, JSON, SOAP bindings or a custom API for non-spatial data.

The OGC service is the second component, which connects the data integration layer to the data analysis layer. It offers a well-known and widely spread standardised interface in proved data structures such as GML, geoTIFF, KML or SVG. In essence, the main difference compared to the non-standard service is that this component provides OGC-standardised data structure over standardised service interface instead of providing a custom interface. The indicated OGC real-time adapter is basically a technological bridge to integrate live data into existing OGC services as existing implementations only support a variety of static data sources such as shape files, ASCII grids, different (geospatial) databases or cascading OGC services (WFS, WMS, WCS). Its implementation has been described earlier in this chapter.

The web service component is the third interface connecting the integration layer with the analysis layer. It is intended for handling non-geographic non-real-time data and for serving it via the http protocol. This component is basically a regular web service, meaning that it implements the request/response communication model, but no pushing mechanism as the non-standard service component does.

The next release of the implementation will enable a wide range of collection and reporting possibilities for integration into existing decision support systems in a variety of application areas, yielding a more complete and accurate real-time view of the city. Furthermore, the integration of Event Stream Processing (ESP) mechanisms will allow for management of pre-defined domain violations and for tracing and analysing spatio-temporal patterns in continuous stream data.

VI. Effects on the Urban Context

From a socio-political viewpoint, Live Geography is primarily targeted at the information needs of local and regional governments. It enables them to respond to the environmental and social challenges of the city and to learn more about the impacts of urban policies and practices. Moreover, it also supports professionals, such as urban and transportation planners, in building or refining their models of urban dynamics. In fact, it can change the work of practitioners that was previously about predicting and accommodating, and which is now becoming more observing and improving. Indeed, this new ability to render all kinds of „machine readable" environments not only provide new views on the city and its environment, but also supply urban and transportation engineers and planners with indicators to evaluate their interventions. For instance, Dan Hill and Duncan Wilson foresee the ability to tune buildings and cities through pre and post occupancy evaluations. They speculate that the future of environmental information will be part of the fabric of buildings [20]. However, this integration opens all sorts of issues regarding sampling, density, standardisation, quality control, power control, access to data, and update frequency.

A complete picture might be hard to achieve with incomplete environmental data patched together by data mining, filtering and visualisation algorithms. Environmental monitoring in the urban context is limited to classic technical issues related to data resolution and heterogeneity. Even mobile sensors do not yet provide high-density sampling coverage over a wide area, limiting research to sense what is technically possible to sense with economical and social constraints. One set of solutions rely on the calibration of mathematical models with only a few sensors nodes and complementing data sources to create a set of spatial indicators. Another, approach aims at revealing instead of hiding the incompleteness of the data. Visualising the uncertainty of spatial data is a recurrent theme in cartography and information visualisation [21]. These visualisation techniques present data in such a manner that users are made aware of the degree of uncertainty so as to allow for more informed analyses and decisions. It is a strategy to promote the user appropriation of the information with an awareness of its limitations [22]. Without these strategies to handle the fluctuating quality of the data, their partial coverage could impact people's perception of the environment, by providing a quasi-objective but inaccurate angle of the content, and potentially „negatively" influencing their behaviour.

Another way to improve the coverage of environment data is to alter the current model whereby civic government would act as sole data-gatherer and decision-maker by empowering everyday citizen to monitor the environment with sensor-enabled mobile devices. Recently providers of geographic and urban data have learned the value of people-centric sensing to improve their services and from the activities of their customers. For instance the body of knowledge on a city's road conditions and real-time road traffic network information thrive on the crowd-sourcing of geo-data the owners of TomTom system and mobile phone operators customers generate. Similarly, the users of Google MyMaps have contributed, without their awareness, to the production the massive database necessary for the development of the location-based version of the application. However, this people-centric approach to gather data raise legitimate privacy, data integrity and accuracy concerns. These issues can be handled with a mix of policy definition, local processing, verification and privacy preserving data mining techniques [23]. These technical solutions necessitate a richer discussion beyond the academic domain on these observing technologies' social implications.

Similar crowd-sourcing strategies have been considered for environmental monitoring with individuals acting as sensor nodes and coming together with other people in order to form sensor networks. Several research projects explore a wide range of novel physical sensors attached to mobile devices empowering everyday non-experts to collect and share air quality data measured with sensor-enabled mobile devices. For instance, Ergo [24] is a simple SMS system that allows anyone with a mobile phone to quickly and easily explore, query, and learn about local air quality on-the-go with their mobile phone. With these tools, citizens augment their role, becoming agents of change by uncovering, visualising, and sharing real-time air quality measurements from their own everyday urban life. This „citizen science" approach [25] creates value information for researchers of data generated by people going on their daily life, often based on explicit and participatory sensing actions. By turning mobile phones [26], watches [27] or bikes [28] into sensing devices, the researchers hope that public understandings of science and environmental issues will be improved and can have access to larger and more detailed data sets. This access to environmental data of the city also becomes a tool to raise the citizen awareness of the state of the environment.

These data gathering possibilities also imply that we are at the end of the ephemeral; in some ways we will be able to replay the city. In contrast we are also ahead of conflicts to reveal or hide unwanted evidences, when new data can be used to the detriment of some stakeholder and policy makers. Indeed, the capacity to collect and disseminate reconfigure sensor data influence political networks, focussing on environmental data as products or objects that can be used for future political action. Therefore, openness, quality, trust and confidence in the data will also be subject of debate (e.g. bias to have people record their observations, who gets to report data and who not). This people-centric view of measuring, sharing, and discussing our environment might increase agencies' and decision makers' understanding of a community's claims, potentially increasing public trust in the information provided by a Live Geography approach.

This raises a real challenge: how can we encourage, promote and motivate environmentally sustainable behaviours on the basis of Live Geography information? Recent work [29] shows how real-time information about cities and their patterns of use, visualised in new ways, and made available locally on-demand in ways that people can act upon, may make an important contribution to sustainability. However, the communication of data collected from pervasive sensor networks may not trigger sufficient motivation for people to change their habits towards a more environmentally sustainable lifestyle. Indeed, this objective of improving the environmental sustainability of a city calls for behaviour modification. It can be induced by intervening in moments of local decision-making and by providing people with new rewards and new motivations for desirable behaviours [30]. These kinds of strategies have been common, for instance, in health and fitness applications. However, when we think about persuasion in the real of environment sustainability, we might want to persuade people of the ways, in which their interests are aligned with those of others [31]. Therefore, this process of alignment and mobilisation, by which one can start to find one's own interests as being congruent with those of others will be critical in the success of these strategies based on Live Geography.

All of the above examples consider the use of sensed data to satisfy public or private requirements of some sort. However, terms like „air quality" are effectively only a surrogate for the health effects of pollutants on people or structures, and there is much disagreement on the public information regulations and whether they are effective. One potential alternative is the direct sensing of structural, or even human, impacts. In other words, people need to be sensitised to their location and their environment, and this information will have to be presented within different contexts, such as public safety, energy consumption or sustainability in order to open new possibilities for exploring the city. In sum, feedback of „sensed" data to the public – either directly or after processing – can potentially change people's perception of the city itself. For example, weather or pollution prediction on a very local scale in space and time is reasonably feasible. Continuous availability of this information could lead to a change in people's individual behaviour by giving real-time short-term decision support. The Live Geography approach can play a significant role in achieving this seminal vision.

VII. Conclusion

Ubiquitous and continuous environmental monitoring is a multi-dimensional challenge, and this is particularly true in the urban context. In this paper we have shown which issues have to be considered for environmental monitoring systems in the city, and have outlined how the *Live Geography* approach can meet these requirements.

It stands for the combination of live measurement data with historic data sources in an open standards based framework using server-side processing mechanisms. Its basic aim is twofold: first, automating GIS analysis processes that currently require a considerable amount of manual input, using new server-based processing tools by providing a wholly standardised workflow; second to replace monolithic measurement systems with an open standards-based infrastructure.

The implementation of the approach comprises the following components: firstly, an embedded measurement device providing sensor data via the standardised OGC Sensor Observation Service (SOS) interface; secondly, a special sensor fusion mechanism to harmonise these data, in our case a GeoServer custom data store. The component provides live measurements in well-established standardised formats (KML, GML, geoRSS etc.). This enables simple integration into specialised GIS analysis software.

A crucial implementation component is the Complex Event Processing (CEP) module, which serves for detecting different patterns – i.e. „events" – in measurement data (threshold exceedance, geo-fencing, etc.), and for quality assurance. The module creates alerts, which are sent via the OGC Sensor Alert Service (SAS) interface. According to the alert, different actions can be taken. Either a message (SMS, email etc.) is sent to subscribers of the alert, or geo-analysis operations is triggered, e.g. a flow model calculation in case of threshold exceedance of local rivers.

Performance of these server-based analysis tasks is satisfactory for near real-time decision support. For instance, a standard Kriging interpolation of 14 sensors takes about 6.4 seconds with an output resolution of 50m using an area of interest of 7x5km, using ArcGIS Server version 9.3. It shall be noted that the largest part (approximately 80%) of the delay is due to sensor response times.

To prove the system's portability, we deployed the same underlying framework in different application areas: environmental monitoring (air temperature variation assessment), public health (air quality and associated health effects) and patient surveillance (monitoring of biometric parameters). The next practical realisation will be an urban „quality of life" monitoring system in order to gain a city-wide picture of spatial and temporal variations of environmental parameters such as particulate matter, pollen concentrations or CO pollution. Their integration with static GIS data (census, traffic emergence, living vs. working places etc.) will maximise significance for local and regional governments as well as for citizens by applying complex GIS algorithms such as kriging or co-kriging to reveal unseen causal correlations between spatial parameters. Other scheduled implementations comprise radiation monitoring and water quality assessment infrastructures. To achieve enhanced usability on the end user side, we are just developing a mobile GIS application, which allows the user to interact with the system while offering a broad range of geographic analysis tools.

Since interoperable open systems in general are not trivial to implement, this motivation has to be initiated through legal directives like INSPIRE, GMES, SEIS and GEOSS. As all of our *Live Geography* implementations are operated in close cooperation with local or regional governments and thematic actors, we think that the Live Geography approach will raise awareness of ubiquitous sensing systems and perhaps trigger profound rethinking process in collaboration and cooperation efforts between different authorities in the city.

Concluding, it shall be stated that the trend towards extensive availability of measurement data requires a paradigm shift in the global GIS market and its applications for urban environmental monitoring. This applies especially to open data accessibility and intensified collaboration efforts. To achieve far-reaching adoption, the establishment of ubiquitous sensing infrastructures will require a long process of sensitising individuals to their spatial and social contexts, and to how to connect local environmental questions to well-known urban issues such as public safety, energy efficiency, or social interaction. In effect, creating a meaningful context around densely available sensor data and distributing specific information layers makes the environment more understandable to the city management, to the citizens and to researchers by „making the abstract real“, i.e. by revealing hidden connections in real-time.

Acknowledgment

Our approach requires expertise in a wide variety of research areas such as sensor networks, data integration, GIS data and analysis, visualisation techniques etc. We would like to thank all contributing groups at the Research Studio iSPACE and at MIT for their valuable inputs and contributions in different stages of the project.

References

[1] Resch, B., Mittlboeck, M., Girardin, F., Britter, R. and Ratti, C. (2009) Real-time Geo-awareness - Sensor Data Integration for Environmental Monitoring in the City. IN: Proceedings of the IARIA International Conference on Advanced Geographic Information Systems & Web Services – GEOWS2009, 1-7 February 2009, Cancun, Mexico, pp. 92-97.

[2] Paulsen, H. and Riegger, U. (2006). SensorGIS – Geodaten in Echtzeit. In: GIS-Business 8/2006: pp. 17-19, Cologne.

[3] Resch, B., Calabrese, F., Ratti, C. and Biderman, A. (2008) An Approach Towards a Real-time Data Exchange Platform System Architecture. In: Proceedings of the 6th Annual IEEE International Conference on Pervasive Computing and Communications, Hong Kong, 17-21 March 2008.

[4] University of Oklahoma (2009) OKCnet. http://okc.mesonet.org, March 2009. (12 May 2009)

[5] Center for Coastal and Land-Margin Research (2009) CORIE. http://www.ccalmr.ogi.edu/CORIE, June 2009 (14 July 2009)

[6] Kansal, A., Nath, S., Liu, J. and Zhao, F. (2007) SenseWeb: An Infrastructure for Shared Sensing. IEEE Multimedia, 14(4), October-December 2007, pp. 8-13.

[7] Paulsen, H. (2008) PermaSensorGIS – Real-time Permafrost Data. Geoconnexion International Magazine, 02/2008, pp. 36-38.

[8] Murty, R., Mainland, G., Rose, I., Chowdhury, A., Gosain, A., Bers, J. and Welsh, M. (2008) CitySense: A Vision for an Urban-Scale Wireless Networking Testbed. Proceedings of the 2008 IEEE International Conference on Technologies for Homeland Security, Waltham, MA, May 2008.

[9] National Oceanic and Atmospheric Administration (2008) nowCOAST: GIS Mapping Portal to Real-Time Environmental Observations and NOAA Forecasts. http://nowcoast.noaa.gov, September 2008. (15 December 2008)

[10] Sarjakoski, T., Sester, M., Illert, A., Rystedt, B., Nissen, F. and Ruotsalainen, R. (2004) Geospatial Info-mobility Service by Real-time Data-integration and Generalisation. http://gimodig.fgi.fi, 8 November 2004. (22 May 2009)

[11] Rittman, M. (2008) An Introduction to Real-Time Data Integration. http://www.oracle.com/technology/pub/articles/rittman-odi.html, 2008. (22 May 2009)

[12] Sybase Inc. (2008) Real-Time Events Data Integration Software. http://www.sybase.com/products/dataintegration/realtimeevents, 2008. (22 December 2008)

[13] Rahm, E., Thor, A. and Aumueller D. (2007) Dynamic Fusion of Web Data. XSym 2007, Vienna, Austria, pp.14-16.

[14] Riva, O. and Borcea, C. (2007) The Urbanet Revolution: Sensor Power to the People!. IEEE Pervasive Computing, 6(2), pp. 41-49, April-June 2007.

[15] Resch, B., Schmidt, D. und Blaschke, T. (2007) Enabling Geographic Situational Awareness in Emergency Management. In: Proceedings of the 2nd Geospatial Integration for Public Safety Conference, New Orleans, Louisiana, US, 15-17 April 2007.

[16] Botts, M., Percivall, G., Reed, C. and Davidson, J. (Eds.) (2007a) OGC® Sensor Web Enablement: Overview and High Level Architecture. http://www.opengeospatial.org, OpenGIS White Paper OGC 07-165, Version 3, 28 December 2007. (17 August 2009)

[17] Resch, B., Mittlboeck, M., Lipson, S., Welsh, M., Bers, J., Britter, R. and Ratti, C. (2009) Urban Sensing Revisited – Common Scents: Towards Standardised Geo-sensor Networks for Public Health Monitoring in the City. In: Proceedings of the 11th International Conference on Computers in Urban Planning and Urban Management – CUPUM2009, Hong Kong, 16-18 June 2009.

[18] Schut, Peter (ed.) (2007) Web Processing Service. http://www.opengeospatial.org, OpenGIS Standard, Version 1.0.0, OGC 05-007r7, 8 June 2007. (19 June 2009)

[19] Bass, T. (2007) What is Complex Event Processing?. http://www.thecepblog.com/what-is-complex-event-processing, 2007. (11 July 2009)

[20] Hill, D. and Wilson, D. (2008) The New Well-tempered Environment: Tuning Buildings and Cities. Pervasive Persuasive Technology and Environmental Sustainability, 2008.

[21] MacEachren, A. M., Robinson, A., Hopper, S. Gardner, S., Murray, R., Gahegan, M. and Hetzler, E. (2005) Visualizing Geospatial Information Uncertainty: What We Know and What We Need to Know. Cartography and Geographic Information Science, 32(3), pp. 139–160, July 2005.

[22] Chalmers, M. and Galani, A. (2004) Seamful Interweaving: Heterogeneity in the Theory and Design of Interactive Systems. DIS'04: Proceedings of the 2004 Conference on Designing Interactive Systems, ACM Press, New York, NY, USA, 2004, pp. 243–252.

[23] Abdelzaher, T., Anokwa, Y., Boda, P., Burke, J., Estrin, D., Guibas, L., Kansal, A., Madden, S. and Reich, J. (2007) Mobiscopes for Human Spaces. IEEE Pervasive Computing, 6(2), pp. 20–29, 2007.

[24] Paulos, E. (2008) Urban Atmospheres - Proactive Archeology of Our Urban Landscapes and Emerging Technology. http://www.urban-atmospheres.net, 2008. (18 June 2009)

[25] Paulos, E., Honicky, R. and Hooker, B. (2008) Citizen Science: Enabling Participatory Urbanism. In: Handbook of Research on Urban Informatics: The Practice and Promise of the Real-Time City. 506 pp., ISBN 9781605661520, IGI Global, Hershey PA, USA, 2009.

[26] Nokia Research (2008) SensorPlanet. http://www.sensorplanet.org, 12 March 2008. (18 August 2009)

[27] La Montre Verte (2009) La Montre Verte City Pulse. http://www.lamontreverte.org, July 2009. (10 August 2009)

[28] Campbell, A.T., Eisenman, S.B., Lane, N.D., Miluzzo, E. and Peterson, R.A. (2006) People-centric Urban Sensing. WICON'06: Proceedings of the 2nd Annual International Workshop on Wireless Internet, ACM, New York, NY, USA, 2006.

[29] Calabrese, F., Kloeckl, K. and Ratti, C. (2008) WikiCity: Real-time Location-sensitive Tools for the City. In: Handbook of Research on Urban Informatics: The Practice and Promise of the Real-Time City. 506 pp., ISBN 9781605661520, IGI Global, Hershey PA, USA, 2008.

[30] Fogg, B.J. (2003) Persuasive Technology: Using Computers to Change What We Think and Do. 312 pp., ISBN 978-1558606432, Morgan Kaufmann, Amsterdam, The Neatherlands, 2003.

[31] Dourish, P. (2008) Points of Persuasion: Strategic Essentialism and Environmental Sustainability. Proceedings of the Sixth International Conference on Pervasive Computing, Sydney, Australia, 19-22 May 2008.