Geoinformatica
DOI 10.1007/s10707-016-0273-y

A semantic HTML based approach for geosensor media

Ruowei Xiao [1] **· Zhanwei Wu** [2] **· Kazunori Sugiura** [1]

Received: 1 February 2016 / Revised: 13 July 2016Accepted: 22 September 2016

Abstract Over the past decade, with rapid development of sensor built-in mobile devices, wireless ad-hoc sensor networks and open sensor platforms, the traditional technocratic Geosensor webs now are experiencing a transition towards an emerging end-user-oriented Geosensor media, through which ordinary users are able to freely process, recreate and disseminate sensor data and information for self-expression, collaborative tasks or collective wisdom etc. In this paper, we discussed the concept, implication and potential issues of Geosensor media. A general framework involving semantic interaction, geospatial interface and participatory sensing was brought out, and a typical prototype was implemented and evaluated.

Keywords Geosensor network · End user service composition · Semantic HTML · Spatial interface

1 Introduction

Geosensor network generally refers to sensor networks consisting of static or mobile sensing nodes that detect, monitor and process events in geographic space [1, 2]. Today, the amount of Internet-connected sensors is proliferating at a phenomenal rate. As a linking dimension, sensor location is generally considered as a key piece of sensor or sensor data information, since every sensor, whether in situ or remote, has a location, and the location of a sensor is highly significant for many applications. The early research of Geosensor network dated back to the early '90s,

✉ Zhanwei Wu
zhanwei_wu@sjtu.edu.cn

Ruowei Xiao
ruowei.xiao@kmd.keio.ac.jp

Kazunori Sugiura
uhyo@kmd.keio.ac.jp

1 Keio University, 4-1-1 Hiyoshi, Yokohama, Kanagawa, Japan

2 Shanghai Jiaotong University, 800 Dongchuan Road, Shanghai, China

Published online: 01 October 2016

when planetary scientists first proposed the concept of standardized description files for sensor location. Subsequent work was done by NASA, the University of Alabama Huntsville and the Committee on Earth Observation Satellites (CEOS), followed by the Open Geospatial Consortium (OGC), which started to promote its Sensor Web Enablement (SWE) activity in 2001 [3].

Besides traditional geospatial technology (GIS, navigation, location services etc), the impact of Geosensor network also radiated to other emerging domains such as the Internet of Things (IoT), Machine-to-Machine (M2 M) communication, optical imaging, Augmented Reality (AR), the smart grid, embedded mobile devices, location-based marketing and social networking. Furthermore, with the pervasion of sensor built-in mobile devices within the last decades, tiny, cheap, geolocated, network-connected sensors have rapidly permeated every aspects of our society, including habitancy, transportation, business and industries. Against a grand background that more and more governments and public institutions data became openly accessible to the public, individuals and social groups are as well encouraged to contribute their data to form a body of collective knowledge. Especially the rising of social sensing technology has greatly expanded the term “sensor” from traditional ICT-enabled devices to a much more widened horizon [4], which further accentuates the potential commercial and social impacts of this research spectrum.

Given the higher popularity and easier accessibility of personal data, many users nowadays are no longer satisfied with standard apps and services provided by manufacturers. People want to manage and analyze data by their own, and present the outcome in a creative way and share with others. The location information derived from the data provided by or about users, on the other hand, are often “fused” with map data and/or computer-aided design (CAD) building data, and also text, image, video or sound in order to generate valuable knowledge and precise insights, which hence pose reciprocal feedback (for example, semantic knowledge) to Geosensor network development in return. Therefore, it is likely for us to infer a turnaround from the traditional technocratic Geosensor web to a novel, generalized, end-user-oriented Geosensor media. In this paper, we define Geosensor Media as a new kind of end-user-oriented system based on Geosensor Web, through which ordinary users are able to freely access, process and disseminate sensor data and related information for self-expression or knowledge communication. The development of this kind of novel systems will expand the application of Geosensor Web technologies from merely engineering domain into broader life scenarios, thus greatly promote the social and business impact of Geosensor technologies. However, from a user’s perspective we have still witnessed a few practical gaps in between:

- *Data Interoperability for both machines and human beings*. The raw physical sensor data usually exists in a format is that easy to be processed by machines, while difficult to understand for ordinary user. For example, a formatted sensor data may look like: <model > ENV </model > <timestamp > 2015–06-05 T04:15:00 < timestamp/ > < humidity_avg > 55.4 </humidity_avg > <temp_avg > 8.1 </temp_avg > <longitude > 144.9404816 </longitude > <latitude > −37.8203537 </latitude>. Users may have to guess what each attribute and their values stand for, but how about “The average humidity and temperature around Docklands Library area were 55.4% and 8.1 °C respectively, at the time 04:15, 05-06-2015”?
- *Sensor Context Reconstruction and Representation.* Without meaningful reference and relations, isolated sensor data does not provide sufficient contextual information to users. A complete context necessarily consists of both physical and social information, which are usually among wide correlations, and comparing sensor data across different districts or

social groups helps to generate overall insights more precisely. Peer sensors are specifically necessary when coming to social sensing. For instance, a single Twitter post reporting earthquake may require collocated and concurrent reports in order to be validated.
- *End User Oriented Tools*. Users currently don't have adequate methods to interact with (their own) sensor data. Although built-in sensors have pervaded our everyday devices like mobile phones, smart watch and etc., still in most occasions users are just receiving the notifications of sensor data unidirectionally and passively. Users have neither access nor methods, to decide which data to be collected, how data to be processed, and in what way results to be shown.

As far as we know, there are plenty of related researches with aims to solve part of issues mentioned above and have already made some progress, for example: Semantic data modeling standards are being developed for web of data where data interoperability is also one of its concern; Participatory sensing and geospatial interface may help to provide richer Sensor Context from information gathering and representation side respectively.

Therefore, in this paper, instead of reinventing these technologies or building an exclusive new system from the bottom up, our purpose is to:

First, propose a general Service Composition framework to better reuse and integrate the state-of-the-art standards and technologies mentioned above, then build a prototype system to verify its feasibility and usability.

Second, since most of these technologies are rather technician and field experts oriented than end user oriented, an extended HTML based approach is introduced to make our Geosensor Media framework, as a whole, more end-user-friendly.

In section II, we briefly introduced research topics related to the issues described above, and both their promises and concerns from a user's point of view. In section III and IV we proposed a general framework for Geosensor Media and provided its typical implementation respectively. User case analysis and related evaluation were presented in section V. And section VI discussed some future issues and close up the whole work.

2 Related work

2.1 Semantic data modeling

Semantic modeling has been widely utilized to interpret Geosensor data in a way close to human thinking and semantic context. Ontological reasoning and other semantic methods are used to tag, classify and analyze raw sensor data in order to extract meaningful knowledge from it. The benefits rest in providing a general method for end users to manage and query over heterogeneous data from varied sensor sources and standards, without actually knowing the internal structures of sensor data [5]. [6] further argued that the inherent semantic links can reveal more valuable tacit patterns than isolated semantic tags. This kind of natural language based abstraction innately complies with human cognition. However, it inevitably results in restricting the numerous activities and complicated patterns, into a limited semantic tags pool at the cost of scalability and/or accuracy. An explicitly fixed application scenario (for example at a campus) or object (for example vehicles), is usually set to narrow down the range of user behaviors and environment information.

To be better adapted to machine processing while still maintain certain level of readability, other modeling languages have also been proposed for semantic abstraction as a compromise solution. In 2002, UC Berkeley and Intel Research proposed a SQL-based interface approach to query sensor nodes over ad hoc sensor networks in a way similar to traditional DBMS [7]. Because of its superiority in processing and retrieving structured data, this SQL-like language and its successors had achieved a considerable user base. Another inspiring example is the Sensor Model Language (SensorML) developed by OGC, extended from standard XML to specifically describe, store and transmit data from sensors, actuators, and their related processes [8]. In this paper, we have adopted SensorML as the bottom-level data model for all Geosensor web resources, and provided a general approach based on extended HTML to establish efficient and extendable semantic interactions between users and SensorML encapsulated data.

2.2 Geospatial interface

Different from semantic interaction, geospatial interface has been leveraged to alleviate the complexity and fuzziness of sensor data in another way, as spatial context often plays a central role of linking dimension that allow users to combine and compare datasets [9]. Geospatial interface based approaches include geographic visualization, Augmented Reality (AR), satellite imagery and etc., with main purposes of either improving the data representation and integration layout, or providing what-you-see-is-what-you-get interaction to manipulate Geosensor data. Given the unprecedented data volume and unprecedented level of data heterogeneity, common data sets that constitute a typical analytic case may actually contain object sensor data (trajectory, heart rate etc), environment sensor data (humidity, temperature, pressure etc), and social sensor data (tax index, crime rate, average housing price etc), which probably differ from spatial grids, sampling rates, measurements, all the way to data structure. Take geographical visualization for instance, it particularly fuses all the information into one single view, generating a more precise panorama across varied devices and fields.

Meanwhile, researchers also pointed out a reality that most of these sophisticated techniques and tools were intentionally designed for specifically trained analysts or data experts, rather than more generic “end users” [10], which leads us back to the third problem mentioned in Section I. Therefore in this paper, we focused more on geographical visualization as a flexible interaction component rather than the technology per se [11], and presented its potentials to collaborate with other methods to constitute an integral framework of high usability.

2.3 Participatory sensing

Contrary to spatial interface, participatory sensing, on the other hand, underlines the importance of social aspects of contextual information. Its rising, mainly due to our everyday usage of sensor built-in devices, enables anytime and anywhere data monitoring of user activities and environments [12, 13], helping address societal issues varied from transportation, public welfare, wellness, ecological environment to education [14]. The significance is that it contributes to collective knowledge across users in different social granularity, ranging from single individuals to groups and communities [13], and most importantly, as [14] put it: “Participatory sensing begins and ends with people.” whose core value lies in the proliferation and rumination of social knowledge. One example is geotagging or geoblogging: In online mapping platforms like Google Maps, Virtual Earth, or Wikimapia, users are allowed to add their own geographic information to web-based displays or modify the contributions of others [12]. By this way, a user is both the producer and the consumer of knowledge, the same way as being both an individual and a member of communities. The main concern for participatory sensing is the privacy issue, as concluded by many previous literature works [4, 9, 13, 15].

To attract as many users as possible, we hence constructed our prototype in an autonomous manner, leaving users with plenty of applicable leeway and full control over data processing. An access control mechanism and safety links were used to ensure that user generated data shall only be shared among authenticated groups that users designated.

3 Geosensor media as a general framework

Though Semantic data modeling, geospatial interface and participatory sensing technology are in the rapid development, and approaching our Geosensor Media vision from different perspectives, still a general system framework is necessary to integrate existed web standard stack and state-of-the-art technologies. Besides, to satisfy the user oriented requirements of Geosensor Media to better process heterogeneous sensor data, generate knowledge and communicate opinions, this framework is expected to be capable of: 1) Shielding users from all the of-bottom details and data heterogeneity, by providing universal data interfaces and customizable control primitives to support user-tailored data processing; 2) Integrating dynamic, variegated sensor data and between multimedia contents to generate rich data presentation; 3) Creating safe, collaborative environment for collecting and disseminating viewpoints and knowledge among massive audiences.

To achieve these features, we hence proposed an End-User Service Composition (EUSC) framework to assemble all the modules, including:

- A SensorML based data interface to provide a unified method of encapsulating and accessing heterogeneous data.
- An extended-HTML-based semantic interaction mechanism for technical or non-technical users, to lower down the learning cost.
- A participatory service composition editor for end users to collaboratively process sensor data, append information and spread knowledge.

We will briefly introduce this system framework in the remaining of this section and provide a feasible implementation adopting this very framework in the next section.

3.1 The EUSC framework for geosensor media

Fueled by the rising of SOA, Web 2.0 philosophy and RESTful architectural style, the entire web environment itself is more likely to be considered as an indispensable development context for most of current sensor web applications and IoT services. As the concept "web-as-platform" indicated, it has become possible nowadays to utilize distributed, cross-domain sensor resources and web services, e.g. sensor data cloud storage, online data visualization tools etc., and further tailor them into highly individualized, situation-oriented applications. Thus, a Geosensor Media system can actually be regarded as an EUSC platform, with the purpose to enable users to freely define their own service composition leveraging existed web services.

According to previous researches, service composition is defined as the process of creating services at runtime [16], while EUSC is defined as service composition where the user who composes a composite service is the same person who ultimately uses it [17]. The life-cycle of EUSC is usually split into four stages: Request, Discovery, Composition and Execution [18]. This linear process, however, is not so consistent with the Geosensor Media framework, since our editing paradigm is much closer to a graph of service nodes, though a basic workflow is presented.

Each service node may achieve certain functionality, and can be connected with each other to solve complex tasks. There exists no isolated execution stage, and once the service node graph is properly constructed, the new content will automatically be generated and visualized. From this perspective, we categorize the Geosensor Media framework into three layers, as shown in Fig. 1:

- Resource Layer, which contains various service nodes to retrieve sensor data or other resources over the web, encapsulates them into a general data model such as: SensorML, and provides universal handlers for all the resources.
- Process Layer, which contains necessary functionality to further process resources, for example: data fusion, calculation, resource transformation, etc.
- Interface Layer, which contains services to translate sensor data and other resources into visual components and ultimately integrates them into a geospatial interface.

Since Geosensor network is actually comprised of Geosensor nodes in a global scale and also other bottom-layer hardware, some of its peculiar physical properties must be taken into account. To further illustrate within this domain, we will discuss more details about this 3-layer framework and further explain its uniqueness from existed technologies in the coming sections.

3.2 Resource layer

Unlike that traditional SOA usually requires small trusted networks of organization applications and firm understanding of complex technology, data schema and programming specifications

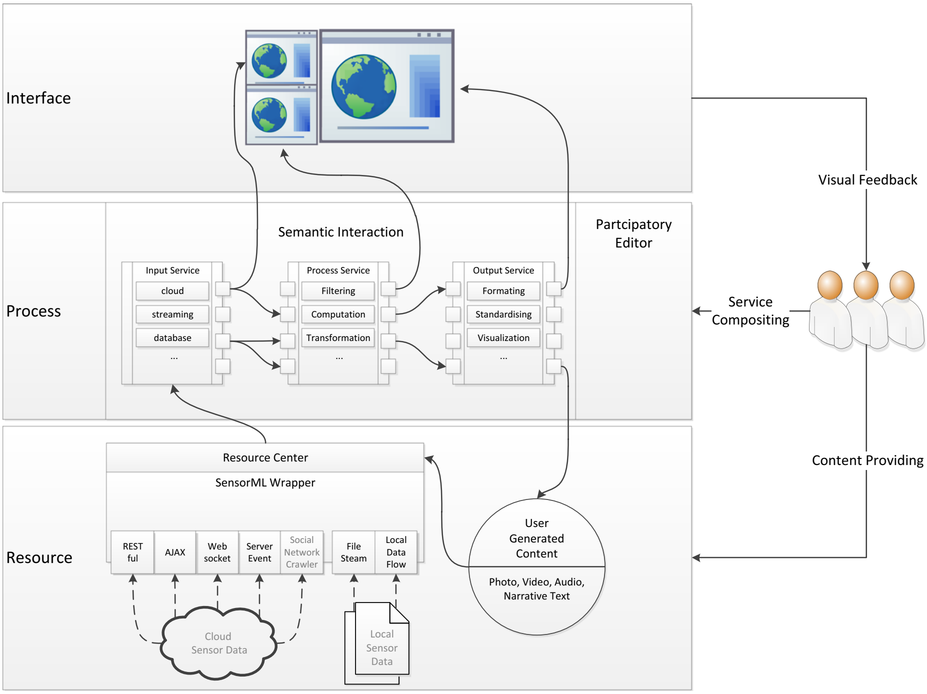


Fig. 1 General framework for geosensor media

[19], the Geosensor Media framework attempts to cope with massive Geosensor nodes spreading throughout the Web, which may differ from device manufacturers, standards, storage, network topology, all the way to data structures and descriptions. Furthermore, proprietary applications and services that handle and process sensor data, e.g. online data analysis and visualization tools, often provide with complicated APIs that require high expertise. These factors have greatly hindered a wider application of Geosensor network into an end user oriented context.

In data modeling and description area alone, research communities and international standards organizations have already come up with rich results: RDF data model together its extensive syntax and grammar like SPARQL, Turtle; JSON extended JSON-LD for linked data; And specifically for sensor web, we have witnessed SensorML proposed by OGC and SSN ontology proposed by W3C. To leverage this relatively mature technology, a standard data model (in our prototype system, SensorML is selected and adapted to framework requirements) should be introduced into the resource layer of Geosensor Media framework, to wrap up Geosensor resources, thus providing unified accessibility to data services in the upper process layer.

Not only Geosensor data, but all the external services to handle them, are supposed to be abstracted and encapsulated as easy-to-use “blocks”, either interactive graphical icons or customizable semantic elements. In this way, the of-bottom details and internal mechanism are managed to be concealed from end users. These heterogeneous resources are identified and exposed in the form of Uniform Resource Identifiers (URI) or Internationalized Resource Indicators (IRI). Being transparent to end users, a resource center or an agent instead, will analyze the data address and type, search if the corresponding “plug-ins” or “pipelines” have been registered, then establish a data channel with the specified sensor or service sources.

3.3 Process layer

The Process Layer is the “control center” where user interactions will be translated into corresponding service compositions. The user defined service composition will then automatically retrieve, process and represent necessary resources to generate new contents which are supposed to meet users’ requirement. Compared with traditional programming method, basic interaction paradigms to lower down the user workload are either Graphic Elements (GE) based, or Domain Specific Language (DSL) based.

The former utilizes graphical icons to represent interactive “blocks” and encapsulate application domain functions and information specific functions within those components, and give a visual face to underlying web resources. GE based composition tools allow end users without adequate programming skills to create customized composition via simple and intuitive user interactions as mentioned. On the other hand, DSL based composition tools provide users with a terse programming language with common syntax and semantics regarding a particular domain [20], in which the interactive “blocks” are often very simple but reusable, shareable semantic entity combinations instead. The tradeoff between kick-start barriers and expressiveness has hence facilitated the hybrid paradigm of GE and DSL [21].

In this paper, we argued that adopting a DSL of larger end user base can better balance between the expressiveness and flexibility of service integration tools, together with usage consistency and learnability for non-technical users [22]. A semantic HTML based method is hence proposed to evaluate our idea, more details will be given in the following sections. The reason why HTML was selected as our primitive DSL is mainly due to: 1) HTML is widely accepted not only by technical community, but also within design and business domains, which results in relatively lower learning barriers and a potentially larger user group. 2)

HTML, as a fundamental component of current web technology stack, provides higher compatibility when integrated with most state-of-art web technologies. 3) HTML shares a consistent syntax system with XML, while the later one is usually used as the description language of many current semantic sensor web standards, include: SensorML, SSN, etc.

And for users who lacks of basic knowledge about HTML grammar, we consider that "programming by demonstration" method can significantly relieve beginners' learning workload and allow them to complete a simple composition task within a few steps.

Further, to support collaborative tasks, a participatory editing mechanism is also necessary for distributed working members for Geosensor Media. In our proposed framework, a participatory editor is tightly coupled with semantic interaction module and allows users to: 1) Retrieve, query over, filter or fuse Geosensor data from multiple sources relying on geospatial attribute and its intercorrelations. 2) Enrich raw data expression by attaching semantic descriptions and multimedia tags, and further extend Geosensor data into a new kind of media contents. 3) Share the data and content to other authorized users or groups and co-process in a realtime manner. 4) Output edited data and content to varied terminals, for example mobile phones, digital signage or wearable devices etc., or pass over to be integrated into other applications or services.

3.4 Interface layer

As depicted in Section 1, we have underlined several notable factors between end users and Geosensor network, among which the representation of sensor data plays an important role that helps users to better understand, manipulate and share Geosensor media content. Compared with ordinary EUSC systems, the data intensiveness of Geosensor network consequently requires a richer data representation method when constructing the application interface.

However, in real cases, integrating multiple kinds of sensors in a visualization interface is still challenging. Massive data volume often overdrives interface update capability, and data heterogeneity (including sampling rate, measurements etc.) makes it impossible to find an all-round solution, leaving alone the expensive learning cost for ordinary users to master related programming skills or software.

Moreover, in certain resource-constraint Geosensor web systems and standalone environments that lack sufficient computing resources or Internet access, it becomes very difficult to realize large-scale data geo-visualization. Therefore, we proposed a bistratal composite structure that consists of both a web map layer and a SVG layer for the interface layer. Since SVG is a platform-independent standard web technology and can be stored, reused and shared among distributed computing systems. And it can as well support data visualization in a standalone environment.

4 A typical implementation: RELOC

To present and evaluate the feasibility of this system framework under current technical condition, we have designed and created our prototype RELOC. Although we limited our prototype to current mainstream web technologies in consideration of compatibility and usability, it may need specific notice that the proposed system architecture shall actually not be restricted to the given technologies.

4.1 A SensorML based general data Interface

As we mentioned in the previous section, the General Data Interface is supposed to, in the first place, be capable of retrieving data from distributed web sources. Ideally, if all systems on the web agree to employ the same technology to provide their data, we can easily obtain any given type of data via a uniform data interface. However, on the actual sensor web, resource heterogeneity in data model and format always remains a difficult issue that impedes the interoperability and data integration. Though OGC and W3C have brought up and promoted a series of standards based on semantic web technology such as SensorML and SSN, still most mainstream open data platforms, to the best of our knowledge, have not yet adopted these standards up until now.

Moreover, the actual web data transfer protocols spread through a very wide and diverse range, including AJAX, RESTful, WebSocket and Server-sent Event, etc. Consequently, in a EUSC system for sensor web we have to develop different data interfaces in adaption to different platforms, and register them into, for instance, a centric data manager or a service agent. Thus, we can introduce heterogeneous resources into Geosensor Media through same kind of data pipelines alike.

In our practice, we developed corresponding data handler for each different sensor data source including local formatted files (.xml, .json, .csv) and online open platforms (OpenSensor.io, Socrata, OGC), and these varied resources were further wrapped up into the standard SensorML data model once the data pipeline was established. Figure 2 shown a temperature sensor's data model derived from SensorML, which represents a typical composition of normal sensor web resources. And by this method, a unified data interface is exposed to end users, through which they can access different data sources by simply appointing their proper addresses, regardless of the underlying differences in internal data structure, description paradigm, storage, or APIs.

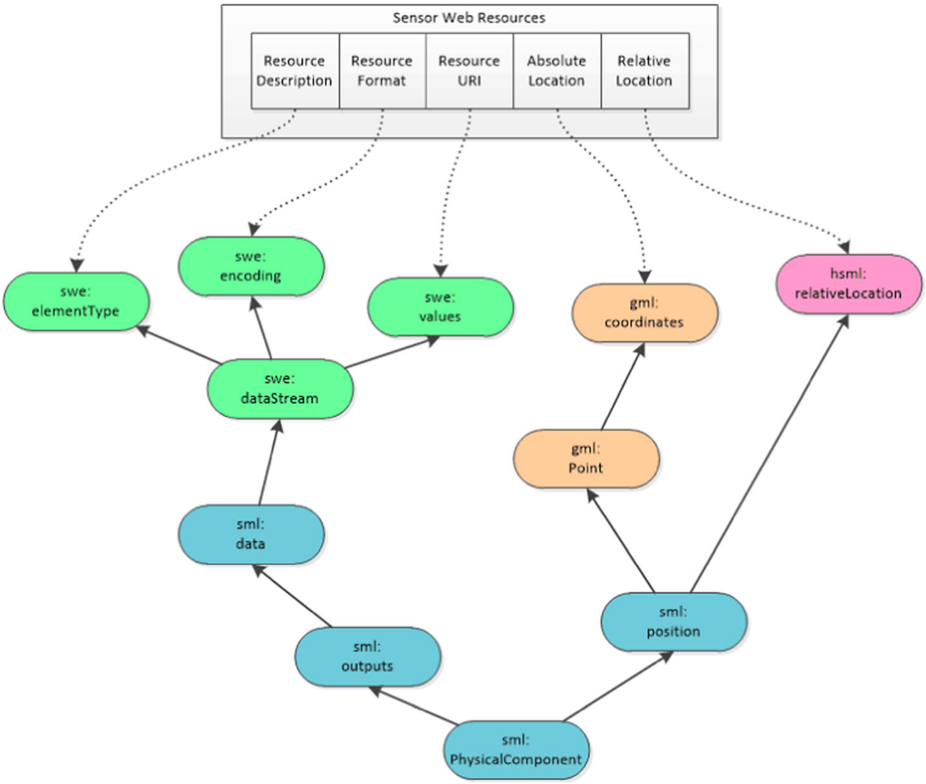


Fig. 2 A data model example of SensorML

4.2 A sensor tag extended HTML for semantic interaction

In this prototype, a Domain Specific Language (DSL) based semantic interaction method is developed. Hyper Text Markup Language (HTML), one of the most pervasive cornerstone web technologies, is selected as the basic DSL mainly due to its innate affinity to human understanding as a semantic description language: compared with general purpose programming languages like C, JAVA, etc., it is obviously one step closer to natural languages.

Therefore, we have extended the traditional HTML and added a handful of semantic tags to specifically process data streams and tackle the interaction between sensor data and context data. Meanwhile by leveraging a series of embedded Cascading Style Sheets (CSS) style syntax within the tags, it is able to freely control and customize visualization outcomes, which, of course, is an optional configuration. And Based on this extended HTML syntax, we constructed our prototype named RELOC.

In RELOC, Geosensor resource is abstracted into a pair of HTML-like tags: <loc > </loc>. In each <loc>, user may use several key value pairs to define its attributes and behaviors, like: <loc key = "value" > </loc > as shown in Fig. 3. A simple example is <loc src = "www.domain-name.com/geosensor"></loc>, in which the HTML "src" attribute is now overloaded to indicate a sensor data source's URI for users to pipe in external resources. Besides, new attributes like "filter", "category" etc. can also be contained by the <loc > tags, in order to meet with specific needs of sensor web service composition.

A < loc > tag is not only able to retrieve data or resources from external sources, but also be connected to other <loc > tags in sequence to accomplish collaborative tasks. All these piped-in resources, together with their handlers and process modules are to form a sequential process flow, to obtain certain purposes, which can be achieved via two specific methods as follows:

First, some services can be encapsulated and introduced into <loc > as specific attributes. For example, users can achieve simple data filtering function by using "filter" attribute, like: <loc src = "" filter = "" > </loc>, where data retrieved from the source will be filtered according to criteria defined in filter attribute. Multiple attributes can be applied simultaneously to the same <loc>, which enables multiple data processing procedures towards same sensor, such as: filtering data and changing units of measurement at the same time.

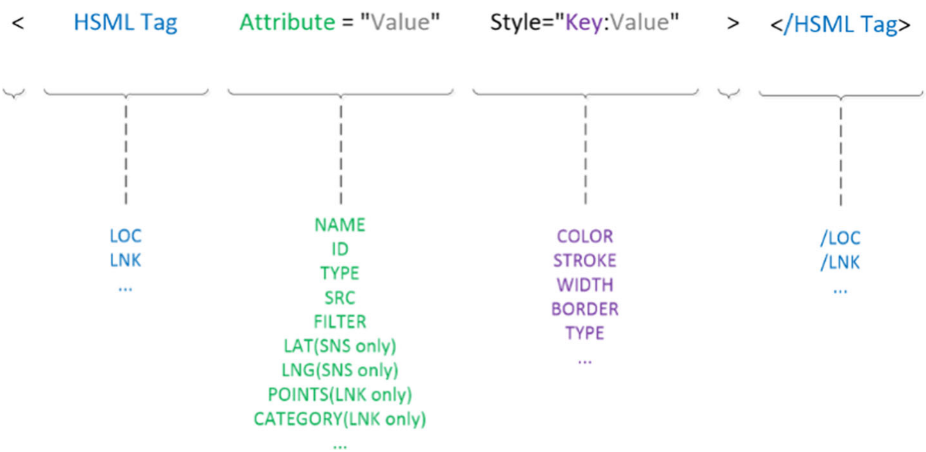


Fig. 3 HSML syntax paradigm

Second, to handle a more complex processing sequence consisting of interconnected <loc > tags, <lnk > tag is introduced. A pair of <lnk > tags represents a linkage between two or more <loc>, usually referring to either a data or control flow. For example, <lnk points = "targetSensor, sourceSensor" category = "identical" > </lnk>, indicates that the "targetSensor" is the equivalent of the "sourceSensor". <lnk > tag is directed, and the first sensor node in the points attribute is regarded as target by default, while remained nodes stands for sources. For example, in <lnk points = "targetSensor, sourceSensor1, sourceSensor2" category = "statisics_mean" > </lnk>, the mean value of sourceSensor1 and sourceSensor2 is calculated and further assigned to the targetSensor.

To notice that, in HSML the <loc > tag does not necessarily refers to a "physical" sensor. It can also be a "virtual" sensor, for example: a web crawler analyzing web data to provide social statistics, or an "agent" sensor that combines and processes data from other sensors to generate new data, so that it can better accommodate general application scenarios.

Moreover, similar to traditional HTML, a < loc > tag can be rendered into a graphical component in visualization stage if proper parameters are set, for instance, an interactive icon on the map. We also provide control over visual elements' appearance following CSS syntax, which we will discuss more details in the coming visualization section. Ordinary HTML contents, including but not limited to plain text, images, audio, video etc., can also be nested inside the <loc > tags to be displayed as additional information window, in this case it will look like: <loc > <div > …</div > </loc > .

By integrating various service modules regarding resource loading, data processing, output transformation and UI representing etc., into HTML style syntax, RELOC demonstrated the feasibility of developing a semantic Interaction for Geosensor Media by simply introducing limited new semantic tags into original HTML. The kind of light-weighted composition pattern of two basic tags with sufficient optional attributes makes HSML highly extendable and flexible, while still maintain the usage consistency and learnability as possible.

4.3 A data-driven geo-visualization Interface

As stated in the previous section, location-based data visualization layer plays a role of linking dimension to combine heterogeneous sensor data together, so that we can grip a clearer panorama of the problem for our decision making. It is particularly important when facing some complicated occasions involving both moving trajectories and dynamic contextual information. For instance, marine animal researchers often combine animal tracking data with environmental information such as sea surface temperature, animal habitat data and etc., to examine and understand habitat utilization and animal behavior in reaction to external forces such as weather or water temperature, which exhibit a different spatial grid and sampling rate. But through the usage of geospatial data visualization, we can bring all data and attributes together into a single view in which a more complete picture of the environment is presented. This enables scientists to identify key locations and times and form new insights into the interactions between the environment and animal behaviors.

Considering the requirement of real-time massive data rendering in many Geosensor related tasks, especially in resource constraint environments, a data-driven visualization method is a necessity. We provided a mapping from RELOC attributes to standard SVG, and a cascaded structure to organize and integrate different visual elements. With each visual element, a data flow is piped in and bound to drive interface update accordingly. The generated SVG layer can be either displayed independently or overlaid on 3rd-party web map layer to form a complete spatial context. At current stage, users are able to pipe in Google map and more data pipelines can be expected in future. If a sensor dataset contains latitude and longitude information, then users can easily pinpoint them visually on map by location attributes.

Visible elements contained in SVG layer, which may be a point, a message box or a marker etc., can be arranged and modified to generate user-customized layouts and appearances. And by restricting interface update operation of a visual element to its minimum scope (for example, only update part of the attributes of a specific element), it is able to update the interface with higher efficiency and lower cost.

4.4 A participatory editor based collaborative GeoSensor media editing flow

With more and more internet-connected, untethered, self-organized sensors coming into place, both the scale of sensor data and the complexity of data tasks have been dramatically escalated to an unprecedented level. It reflects an increasing need for multi-users, cross-platforms and real-time collaborative interactions to support interdisciplinary, multi-granularity data fusing, editing and sharing. RELOC provides a collaborative editor to support users to share the data and content to other authorized users or groups and co-process in a real-time manner. Output edited data and content to varied terminals, for example mobile phones, digital signage or wearable devices etc., or pass over to be integrated into other applications or services.

While privacy has been always a concerning issue within the area, an access control mechanism is involved to ensure that users disclose their sensitive private data (for example GPS, biophysical data etc) only to their trusted friends and communities, with part of or full access (for example, readable, referable, editable and full control) according to the trust levels.

5 Use case

5.1 Lucid manipulation and visualization tool for open sensor data

Nowadays, many government offices have provided massive open data, including various sensor data. However, it’s not so easy for users to understand the meaning of open sensor data by merely look over the original dataset. Data manipulation and visualization may help to better understand the meaning of datasets. However, these tasks often require complex programming skills which usually lacked by ordinary users. To discuss the potential of using RELOC framework to help users deal with open data easier, this use case is designed.

Users can use html-like semantic commands to retrieve data from open data server, filter data and visualize them in a spatial interface, as shown in Fig. 4. The CSS-like grammar is designed to customize data visualization layer. Texts, images or other contents can be embedded in the <loc > tag to provide extra information. Other visualization effects can also be added if needed, such as: heatmap, waveform or some statistic graph. In this use case, open data from the city of Melbourne was used. Since many cities and districts around the world have developed their open platforms based on Socrata or CKAN, RELOC framework allows users to handle most open data manipulation tasks worldwide by utilizing the same mechanism.

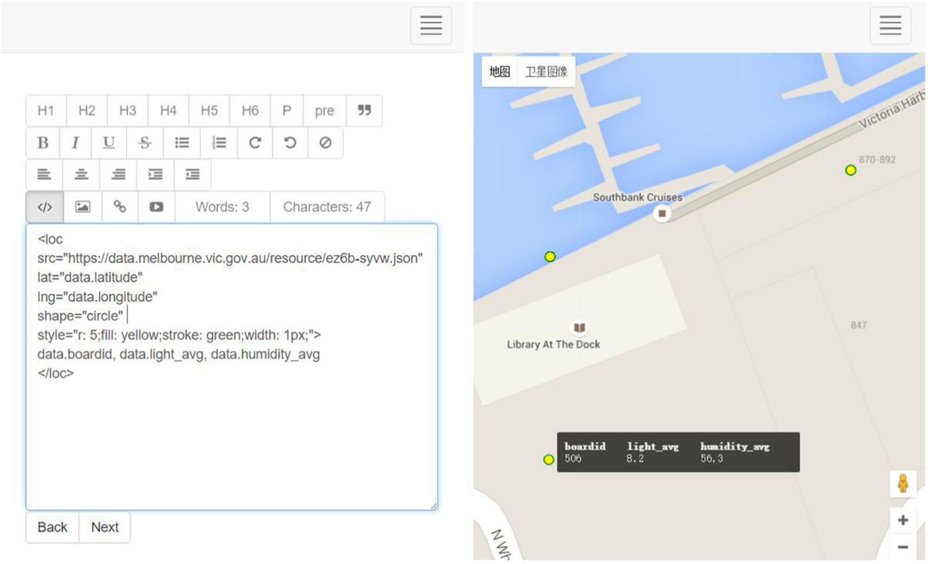


Fig. 4 Use case 1

5.2 Real-time sensor data fusion and visualization

The OGC's recent Sensor Web Enablement (SWE) initiative is attempting to establish a standard for fusing sensor and user data based on spatial context. However, this standard is still in development. The following use case was designed to turn OGC's vision into a practical reality to some extent and provide a feasible example.

Suppose a user is driving his/her car on the street. There maybe exist many sensors around the user, for example: temperature and humidity sensors in the car, or biosensors embedded in the user's wearable device. Fusing and visualizing these sensor data and user generated message can help reconstruct a complete context to better understand the situation. However, because all these sensors may came from different manufacturers, which implies the data may be maintained by different data platforms and interfaces, It becomes difficult for ordinary user to fuse all the heterogeneous data.

As shown in Fig. 5, the designer used two <loc > tag to load real-time sensor data from two different sources: one from OGC's OpenSensorHub, which indicates the current location of the car; one from realtime.opensensors.io, which indicates the real-time heart rate of the driver.

The current car location is visualized as the blue circle with a red border, and the previous ten history data are shown as blue circle with smaller radius according to sequence. The heart rate sensor data is shown in the information window and moving accordingly with the car, as specified in the <lnk > tag. If more sensors added, extra data can also be fused and visualized by semantic commands written in <lnk > tag.

5.3 Participatory sensor data editor

Participatory sensing usually involves multiple users to maintain sensors and manipulate data. To discuss how RELOC framework helps to simplify the situation, we designed this use case.

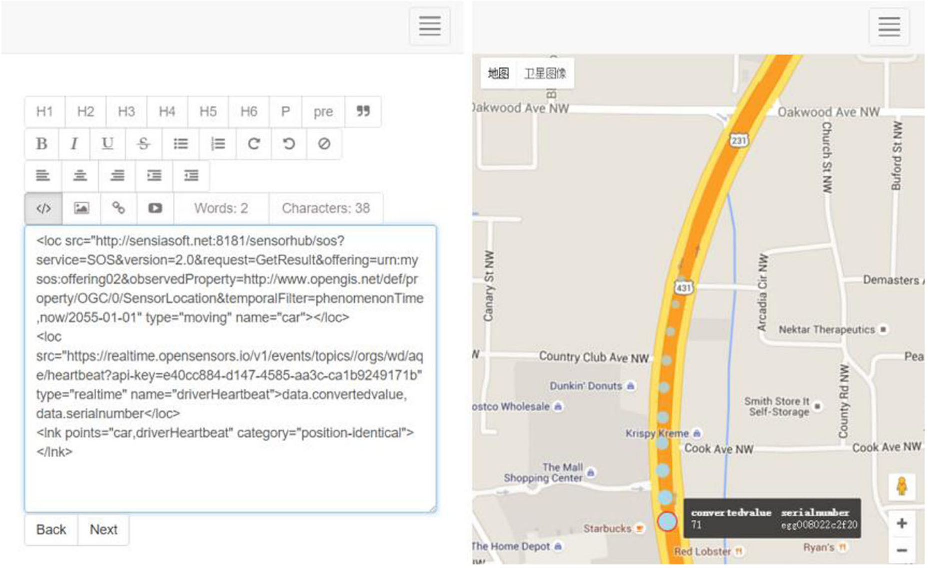


Fig. 5 Use case 2

Suppose two users have established two temperature sensors on two different buildings and created their own web services to publish their data, while the third user is in charge of collating the statistical data. The RELOC framework fulfills these tasks with a participatory editor.

Fig. 6 below shown the third user's view when using participatory editor. While in the same task, semantic commands pushing sensors data have already been added by other users, i.e. the sensor owners, which displayed in gray. The third user can simply edit a < loc > to specify the location and a < lnk > to link sensor data and visualize the statistics, for example: average value.

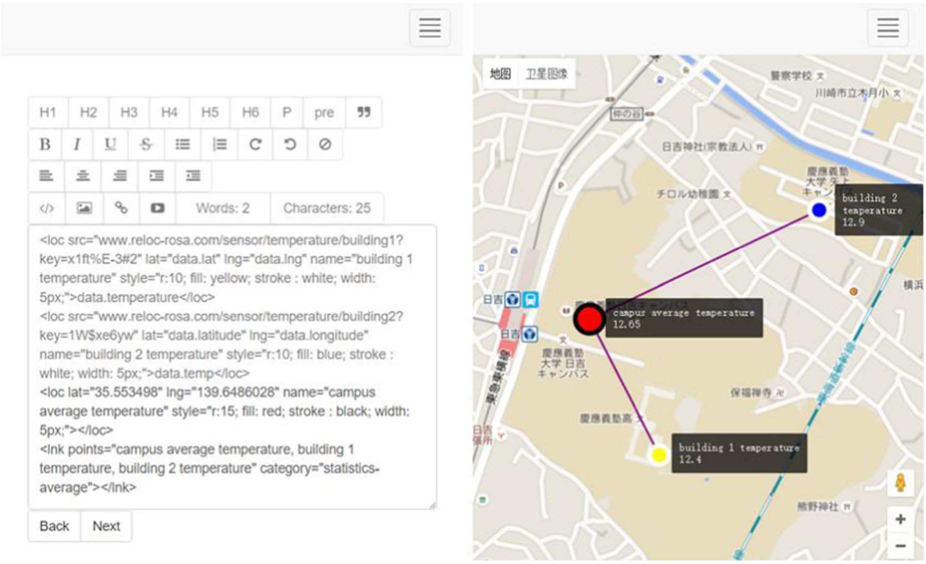


Fig. 6 Use case 3

6 Evaluation

Researchers have already proposed many criteria for information system evaluation. For example, DeLone and McLean's DMSM is one of most cited models which consists of six categories of IS success: system quality, information quality, use, user satisfaction, individual impact and organization impact [23]. Since our prototype is designed to help users interact with sensor naturally, promote sharing and communication of sensor data and user generated content among social group and community to form a body of collective knowledge, evaluation in this research is mainly focus on satisfaction (of users) and impact (of user and social group). More specifically, the impact in this research means the ability of our prototype to help user better contribute and seek information among social groups and community.

Research shown, the easiness to learn and use is a major factor to elevate user satisfaction [24]. While sociability has significant impact on successfulness of social participatory information system, such as: online community or social network, by promoting information contribution and seeking behavior [25]. Thus, in this research, our evaluation mainly focus on learnability and sociability. In learnability part, we studied the speed of learning and first time performance. While in sociability part, we studied whether the prototype system could promote user's willing of sharing and ability of finding.

20 students were recruited as participants for the evaluation experiment. According to previous HTML learning experience, participants were divided into 3 different categories: with no previous HTML learning experience, less than 1 year learning experience, more than 1 year learning experience. All participants had never used prototype system before.

6.1 Learnability

In learnability test, participants were required to learn the functionality of the prototype system by watching a tutorial video. A brochure was provided to explain the usage of RELOC. The prototype system was running on a PC client with connection to server. Student might watch the video as long as they want and test semantic HTML commands on the PC. After the learning process, students were required to execute a relatively easy task to test first time performance. The task included 3 steps: Retrieving data – use RELOC to load data from data source with specified URL; Editing sensor information – add extra description information, such as: location, name and label; Visualizing – tweak the visualization effects by providing style commands. Time consumption of each step was recorded. Mean time were shown in the Fig. 7.

The mean time of learning process had statistical relation with learning experience ($p = -0.8$, $a < 0.01$), which implied the previous HTML learning experiences did help to learning process. However, from the actual value, we could tell that non-experienced users spent only 50 % more time than experienced users on average. This relatively insignificant difference suggested that there existed low learning barriers for non-experienced users.

6.1.1 First-time performance

With the help of the reference brochures, all participants successfully completed all 3 steps of the task. Results shown, previous HTML learning experiences could significantly improve the first-time performance. An experienced user may complete the whole data retrieval, processing and visualizing task in just several minutes without any extra programming. Besides, even for non-experienced users, the performance was acceptable. In the interview after the test, most

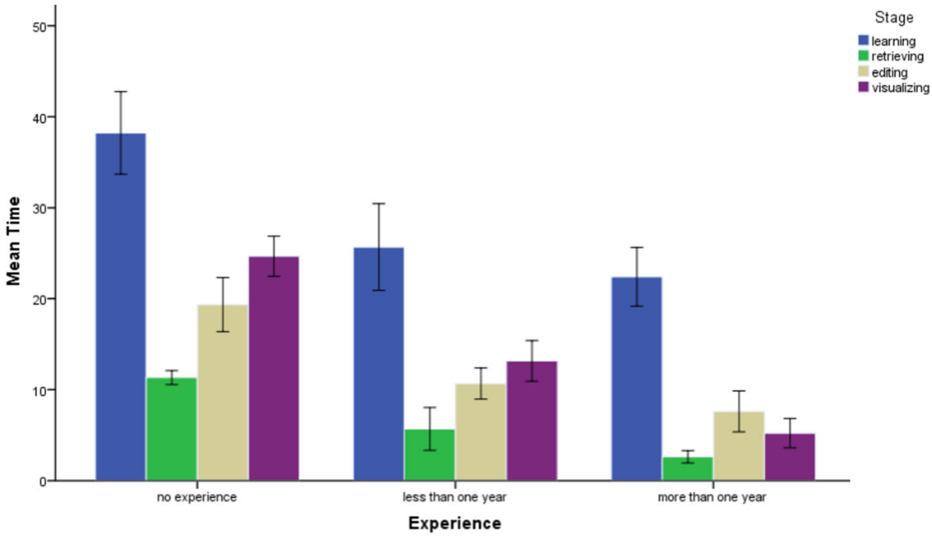


Fig. 7 Learnability test results

participants, either have previous HTML experiences or not, agreed that the semantic interaction used by RELOC was easy to use and very effective.

6.2 Sociability

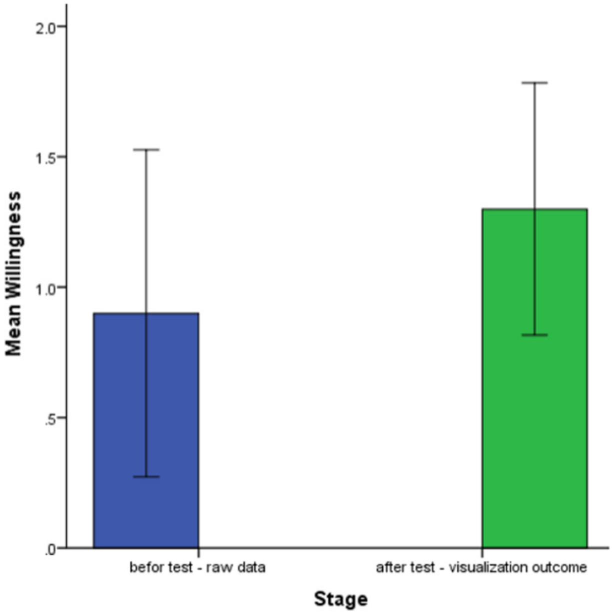
In sociability test, 10 students with best performance in learnability test were invited to design sensor data visualization based on environmental sensors on campus. 100 sets of environmental sensor data, including: PM2.5, temperature, humidity and noise level, are provided. Raw data includes sensor name, position, source URL and simple description provided by their owners. Before the test, students were asked about the willing of sharing raw data to their friends if they were the owner. After the test, a URL that linked to users generated contents (including data visualization and other inserted contents) was create. Willing of sharing this URL to their friends were asked again. Students' willing of sharing were rated as 5-point LIKERT scale: absolutely not (−2), possibly not (−1), not decided (0), possibly yes (1), absolutely (2).

Results shown, the tested students were more willing to share self-defined visualization outcomes than raw data, with a statistically significant difference (sig = 0.021). In the interview after experiment, we asked participants for the reason why they decided to share or not. Answers shown, the sharing decision mainly depended on: Usefulness for others; Sense of accomplishment; Sense of pleasure. These outcomes indicated that, visualized data was more likely to be shared because they were more useful to others, or able to bring more accomplishment and pleasure to owners, as shown in Fig. 8.

6.2.1 Ability of finding

To find specific sensors, a query command usually consists of one or more keywords. If sensor's information is labeled adequately and properly, it can be easily found by matching keywords and labels. However, this ideal situation rarely happens. Most sensors, especially those in open sensor platforms, lack of necessary description to provide clues for searching. Geosensor media may help to address this problem from two aspects: First, the semantic

Fig. 8 Sociability test results



interaction used in Geosensor media will naturally generate more semantic labels to describe sensors; Secondly, analyzing sensor links generated by Geosensor media users will help to establish relation between sensors, thus providing possibilities to design new searching methods that based on relations. To evaluate these effects, we designed an experiment to compare the performance of experimental group (use relation based searching method) and control group (use keywords matching method).

As shown in Fig. 9, three types of links were analyzed: direct link, indirect link and geospatial distance, and used to compute the relation strength between sensors.

Direct links are those links defined explicitly in <lnk>, for example: <lnk points = “sensor0321, sensor0217” > </lnk > defined a direct link between sensor0321 and sensor0217. The strength of direct link, denoted as S_{DL}, is decided by the frequency and category of <lnk > shown in user generated semantic HTML text.

$$S_{DL} = \sum Freq_{\langle lnk \rangle} \times Coef_{\text{category}}$$

Indirect links are those links defined implicitly by semantic labels contained in user generated semantic HTML text, such as: name attributes. For example, sensor0321 and sensor0217 both had the same label “sjtu”, which would generate an indirect link between them. The strength of indirect link, denoted as S_{IL}, is decided by the count of semantic labels.

$$S_{IL} = \sum Freq_{\text{semantic}}$$

Distance means great-circle distance between two sensor’s latitude-longitude points, which is calculated by the Spherical Law of Cosines.

$$\mathrm{D} = \mathrm{acos}(\sin\varphi_1 \times \sin\varphi_2 + \cos\varphi_1 \times \cos\varphi_2 \times \cos\Delta\lambda) \times \mathrm{R}$$

Where φ is latitude, λ is longitude, R is earth’s radius (mean radius is 6371 km).

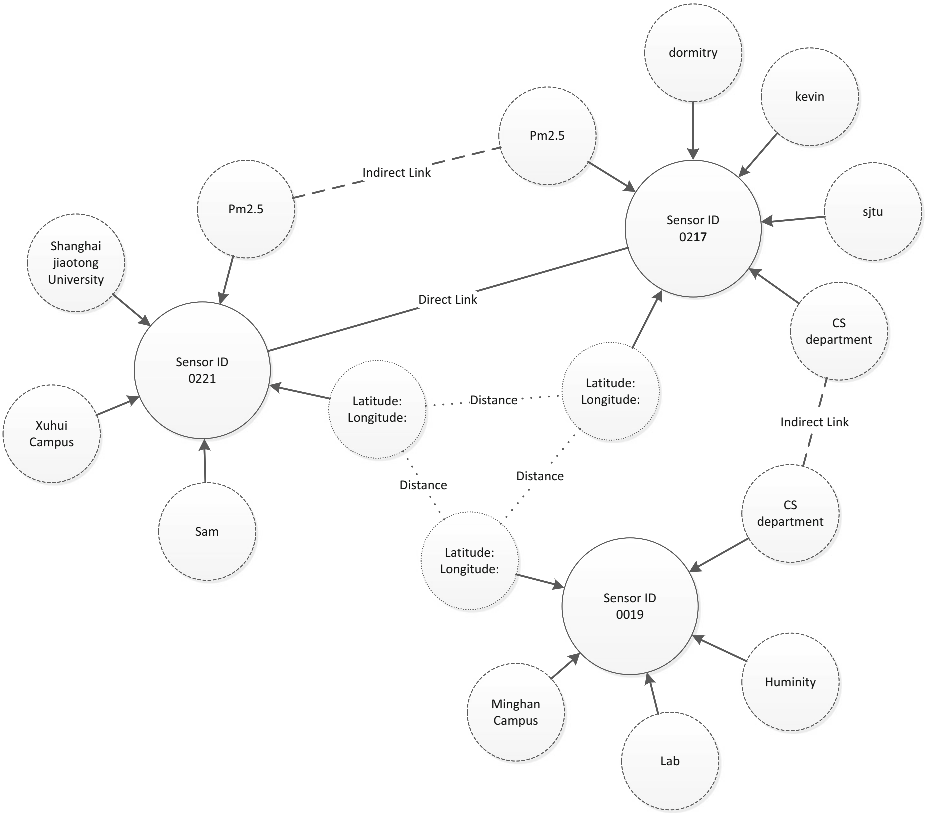


Fig. 9 Linking model used to compute relation strength

Finally, the strength of relation between two sensors, is the weighted sum of S_{DL}, S_{IL} and D:

$$\mathrm{R} = S_{DL} \times w_1 + S_{IL} \times w_2 + D \times w_3$$

Generally, S_{DL} has the most significant weight, because it comes directly from the user's knowledge about the relation between sensors.

After calculating relation strength between each two sensors, a weighted graph contains all the sensors (as nodes) and their relations (as edges) is constructed and used for relation based on searching in experimental group. We simulated a set of search queries, for example: "PM2.5 Minhang campus dorm", executed them in both experimental group and control group. Results shown, the experimental group provided less failure rate (3 %) vs control group (26 %) and returned 87 % more valid records in average. An example of comparing result from experimental group (right) and control group are shown in the Fig. 10. Records marked in green rectangle are valid records that were missed in control group.

The experimental group also generated a few invalid records (marked in red rectangle). However, since we can use the degrees and edge weights in relation graph to sort the records, the invalid records can be easily excluded by users because they usually have less degrees and edge weights.

Fig. 10 Comparison of searching results from relation based method and key word matching method

6.3 Subjective comparison

Finally, we have looked over several representative researches on web based Geosensor information processing and sharing platforms (especially with end user service composition features) and made paralleled comparisons over specific factors related to performance, expressiveness and the degree of user friendly. we have selected eight indexes, they are respectively: service composition style, unified data model, cross-domain data integration, supported data interface, data process functions, geovisualization level, collaborative task, extendability, and learning curves. The results were shown below:

As shown in Table 1, RELOC created a unique user experience by providing end user service composition, geo-visualization and participatory features simultaneously, which can help end users, even without technology background to easily create their sensor based contents collaboratively. We believe this kind of user experience will strengthen the relation between users and sensors, generate more knowledge and insights, and gradually form a new media of users and sensors.

7 Conclusion

In this paper, we argued that a new trend had emerged to better meet the end user's requirement to handle massive sensor resources with very low or even none technical

Table 1 Subjective comparison with paralleled researches

Feature	RELOC	Sensor pedia	Sensor masher	Medusa	TerraFly GeoCloud
End-user service composition	DSL	None	GE	DSL	None
Unified data model	SensorML	Unknown	RDF	Self-Defined	None
Cross-domain data integration	Support	Support	Support	N/A	N/A
Geo-visualization support	SVG/JS	JS	JS	None	JS
Participatory editing	Yes	No	No	Yes	No
User generated content	Rich and integrated	Rich but dispersed	Few	Few and dispersed	None
Extensibility	High	High	Medium	High	Medium
Learning curve	Low to medium	Low to High	Low	Medium	Medium

efforts, which drove traditional technocratic Geosensor webs to become a more user oriented Geosensor Media. To adapt to this new trend, we proposed a general framework that integrates semantic based EUSC, spatial visualization and participatory sensing to interact with spatially distributed sensor in a natural, collaborative way. We believe this framework can give users, even without programming skills, an easy-to-use tool to generate knowledge and insights from massive sensor data, and share with others like traditional media, thus turning existing Geosensor web into a Geosensor media. To embody our vision, we developed a prototype call RELOC with several use cases. The validation experiment partially supported our idea. Users reported that they could easily use this tool to develop their ideas from massive data, collaborate with each other to present visualized knowledge and share with audiences without programming. This makes users more willing to share their data and provide knowledge generate from the data. This process also generates more semantic labels describing sensor resources, hence enhancing computer's understanding about human knowledge regarding sensor data, which in return, will help us to develop a smarter sensor web.

Acknowledgments This work was supported by the Science Foundation of Ministry of Education of China grant 14YJC860029.

References

1. Nittel S (2009) A survey of geosensor networks: advances in dynamic environmental monitoring. J. Sensors 9(7):5664–5678
2. Nittel, S., Labrinidis, A., Stefanidis, A. (2006) Introduction to advances in geosensor networks. In: International conference on GeoSensor Networks, pp. 1–6. Springer, Berlin Heidelberg
3. Why is the OGC involved in Sensor Webs, (2016) http://www.opengeospatial.org/domain/swe
4. Rosi, A., Mamei, M., Zambonelli, F., Dobson, S., Stevenson, G., Ye, J. (2011) Social sensors and pervasive services: Approaches and perspectives. In: 2011 I.E. International Conference on Pervasive Computing and Communications Workshops, pp. 525–530. IEEE Press
5. Ilarri S, Stojanovic D, Ray C (2015) Semantic management of moving objects: A vision towards smart mobility. J. Expert. Syst. Appl. 42(3):1418–1435
6. Sun Y, Yan H, Lu C, Bie R, Zhou Z (2014) Constructing the web of events from raw data in the web of things. J. Mob Inf Syst 10(1):105–125
7. Madden S, Franklin MJ, Hellerstein JM, Hong W (2002) TAG: A tiny aggregation service for ad-hoc sensor networks. J. ACM SIGROPS Operating Systems Review 36(SI):131–146
8. Sensor Model Language(SensorML) Version 2.0, (2016) http://www.opengeospatial.org/standards/sensorml
9. Balazinska M, Deshpande A, Franklin MJ, Gibbons PB, Gray J, Hansen M et al (2007) Data management in the worldwide sensor web. J. IEEE Pervas. Comput. 6(2):30–40
10. Andrienko G, Andrienko N, Dykes J, Fabrikant SI, Wachowicz M (2008) Geovisualization of dynamics, movement and change: key issues and developing approaches in visualization research. J. Inf. Vis 7(3–4):173
11. MacEachren AM, Gahegan M, Pike W, Brewer I, Cai G, Lengerich E, Hardistry F (2004) Geovisualization for knowledge construction and decision support. J. IEEE. Comput Graph 24(1):13–17
12. Elwood S (2008) Geographic Information Science: new geovisualization technologies–emerging questions and linkages with GIScience research. J. Prog. Hum, Geog

13. Lane ND, Miluzzo E, Lu H, Peebles D, Choudhury T, Campbell AT (2010) A survey of mobile phone sensing. J. IEEE Commun Mag 48(9):140–150
14. Goldman J, Shilton K, Burke J, Estrin D, Hansen M, Ramanathan N et al (2009) Participatory Sensing: A citizen-powered approach to illuminating the patterns that shape our world. J. Foresight & Governance Project, White Paper:1–15
15. Khan WZ, Xiang Y, Aalsalem MY, Arshad Q (2013) Mobile phone sensing systems: A survey. J. Commun. Surveys. Tuts. 15(1):402–427
16. Tosic, V., Mennie, D., Pagurek, B. (2000) Dynamic Service Composition and Its Applicability to E-Business Software Systems–The ICARIS Experience. J. Object-Oriented Business Solutions ECOOP
17. Ridge A, O'Neill E (2015) Establishing requirements for end-user service composition tools. J Requir Eng 20(4):435–463
18. Silva, E., Martínez López, J., Ferreira Pires, L., Sinderen, M. (2008) Defining and prototyping a life-cycle for dynamic service composition. J.
19. Anjomshoaa, A., Bader, G., Tjoa, A. M. (2009) Exploiting mashup architecture in business use cases. In: IEEE International conference on network-based information systems. Indianapolis, USA. ISBN (Vol. 1129004680)
20. Maximilien, E. M., Wilkinson, H., Desai, N., Tai, S. A. (2007) Domain-specific language for web apis and services mashups. In: Proceedings of Service-Oriented Computing–ICSOC, 13
21. Maraikar Z, Lazovik A, Arbab F (2008) Building mashups for the enterprise with sabre. Proceedings of Service-Oriented Computing–ICSOC:70–83
22. Aghaee S, Nowak M, Pautasso C (2012) Reusable decision space for mashup tool design. the Proceedings of the 4th ACM SIGCHI symposium on Engineering interactive computing systems: 211–220
23. DeLone WH, McLean ER (1992) Information systems success: the quest for the dependent variable. J. Inform. Syst Res 3(1):60–95
24. Preece J (2001) Sociability and usability in online communities: determining and measuring success. J. Behav Inform Technol 20(5):347–356
25. Phang CW, Kankanhalli A, Sabherwal R (2009) Usability and sociability in online communities: A comparative study of knowledge seeking and contribution. J. J. Assoc. Inf. Syst. 10(10):721

Ruowei Xiao is a Ph. D student at Keio University, Japan.

Zhanwei Wu is an assistant professor at Shanghai Jiaotong University, China.

Kazunori Sugiura is an associate professor at Keio University, Japan.