

HIS-KCWater: Context-Aware Geospatial Data and Service Integration

Jejung Lee
Department of Geosciences
University of Missouri - Kansas City
Kansas City, MO 64110
leej@umkc.edu

Yugyung Lee, Sanket Shah
Department of Computer Science and
Electrical Engineering
University of Missouri – Kansas City
{leeyu, sts4zc}@umkc.edu

James Geller
Department of Computer Science
New Jersey Institute of Technology
Newark, NJ 07102
geller@oak.njit.edu

ABSTRACT

The nature of today's geographically and managerially distributed geospatial information sources makes the interoperability across different sources from different organizations difficult. The integration of real-time water quality assurance data with geographic data is even more challenging. In this paper, we present a context-aware geospatial data and service integration framework that is based on the combination of a syntactic model, a semantic model and a pragmatic model using Semantic Web technologies. This model is context-aware, with the ability to analyze existing dependencies, predict causes and effects and provide context-aware services (which information services are relevant, how to perform the services, how often they are needed, etc). As a proof of concept, we demonstrate the HIS-KCWater system for supporting users in analyzing real time watershed data, predicting the water quality using hydrologic model simulators, interpreting the results, dynamically forecasting problems and generating alerts about water quality issues.

Categories and Subject Descriptors

H.4 [Information Systems Applications]

General Terms

Management, Design

Keywords

Context-aware geospatial service, semantic data service, real time data integration and analysis, water quality service

1. INTRODUCTION

The US Government considers it part of its mission to collect and disseminate information of interest to the general public and to its own federal, state and local agencies about topics such as the weather, population statistics, etc. Nowadays, this dissemination is often done through the World-Wide Web.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

SAC'07, March 11-15, 2007, Seoul, Korea.

Copyright 2007 ACM 1-59593-480-4/07/0003...\$5.00.

There are staggering amounts of geospatial information available on Web pages; even more information lies hidden in the databases behind some of these pages. Those geospatial data are spread out over the Web, with formats and contents depending on the developers who created them. While a surprising amount of relevant data are available on the Web, they come in many different formats and organizations. There are no universal standards for this kind of data, and information is often made available in exactly the format that it was created in. Thus, interoperability across different services from different organizations suffers. The integration of real-time water quality assurance data is even more challenging.

In recognition of the great success of the Web on one hand and its failures due to problems, some of which were outlined above, on the other hand, the vision of a Semantic Web was born [3]. In the Semantic Web vision ontologies play an important role and support semantic index structures, Web services, and semantic integration. Thus, as described in the vision of the Semantic Web, it is our goal to off-load the discovery, processing and integration of Web data from the user onto the system. The Semantic Web should provide what the user wants, when he wants it, making the process of finding and combining the different data and knowledge items completely transparent to him.

2. RELATED WORK

During the enormous growth of the Web in recent years, Web data discovery mechanisms have been developed based on crawling techniques [1]. The process of Web crawling and classification can be significantly improved by using advanced ontological models such as the ontology-based discovery of distributed RDF-based Metadata and documents [9], ontological models describing the contents of Web-pages using text mining technology [10], Metadata [12] and reasoning with OWL [14].

Ontologies play an essential role in the geosciences domain describing correspondences and interrelations among heterogeneous spatial entities [13]. The semantic approach based on the concept of geographic entities enables the seamless integration of geographic information [7] and integration of web search in geographic information systems [15].

Automatic extraction and Metadata encoding of information [11] have been developed using geometry, image interpretation and spatial data mining. [8] integrates applications to extract data from Web resources, enabling the discovery of trends, patterns and relationships. Geographical terminologies using ontologies have been reported in [11]. [2] is collating information from multiple

sources while considering information such as geographic locations. Our framework differs from these efforts by our emphasis on the integration of existing ontologies with each other and with Metadata and the generation of a semantic index structure by crawling the Web in search for hydrologic information.

3. THE WATER QUALITY DOMAIN

3.1 Needs for a Context-Aware Model

Water is one of the most important resources on Earth. Water is abundant in streams, rivers, lakes, groundwater, etc. and is used for drinking, swimming, fishing, commercial purposes, etc. Water is also an essential component in wildlife habitats and aquatic ecosystems. However, water is a most vulnerable resource to the threats of contamination from a variety of human activities. Pollutants such as fertilizers, pesticides, construction debris, and animal waste are often carried by the surface runoff from rainwater, snowmelt, or irrigation discharge into surface water. The interactions between surface water and groundwater through the process of infiltration and base flow may cause secondary contamination of groundwater. Therefore, conservation of water quality is becoming a critical issue for public water suppliers, policy makers and ordinary residents.

Real-time monitoring, especially of Non-Point Source (NPS) pollution, is critical because such sources cannot be clearly identified and delineated in size and location. The US has a broad network of real-time monitoring stations for streams, water quality, groundwater, and precipitation, and produces enormous amounts of data every day. Although the US Geological Survey (USGS) and the Environmental Protection Agency (EPA) provide such data through their websites, the analysis of water quality requires additional processes using professional knowledge and skills. The motivation of the present research is to establish a context-aware water quality monitoring and analysis system to give users easy access to water quality data and analysis results that are automatically produced in real-time.

Research on how to manage the cycles of water resources is still under development. To provide context-aware services for water management, data and services need to be collected from diverse sources and integrated to support decision making. Currently, needed data are manually searched and downloaded based on a specified region of an interactive map. A context-aware automatic discovery capability that can identify a set of data for water management would be of great value for hydrologic researchers.

3.2 Motivating Scenarios

Sustainable management to make water available when and where it is needed is becoming more difficult due to complex climate changes, water pollution, and haphazard use of water. To make sustainable water management possible, a complete understanding of complex hydrologic cycles and water quality issues is necessary. Let us assume that we want to analyze the effect of a sudden reduction in precipitation in Kansas City (Missouri) on the quality of water, the local population and the groundwater

level. Thus, a state planner would need to find both population data and precipitation data. The US Government is keeping excellent sources of hydrologic data available on the Web. For instance, USDA Climate data¹ contains summaries of precipitation for several states, divided into different water resources. A researcher interested in obtaining this kind of data for Kansas City would need to download all the data into a spreadsheet and compute results on his own, taking into consideration borderline areas of Kansas and Missouri. While the necessary information is accessible on the Web, it is distributed over many different Web sites. Another issue is that drinking water quality and safety regulations vary depending on location, time, the condition of the source water (e.g. contaminants) and the treatment of the water.

For analysis purposes, it is necessary to know where drinking water comes from. Water may come from surface sources, such as lakes, rivers, and reservoirs which may be close to the community or many miles away. Second, more importantly, we need to know what contaminants were found in drinking water, the levels at which they were found, and the actual or likely source of each contaminant. It is also important to identify potential sources of contaminants and prevent them from contaminating drinking water. Young children are particularly susceptible to the effects of high levels of certain contaminants, including nitrates and lead. If water sources do not meet government standards, it is necessary to know whether they have potential acute or chronic health effects. In such a case, an alert service is crucial to find an alternate source of water or take precautions, such as boiling the water. In addition, this information should be promptly disseminated through the local media, direct e-mail, phone, or other means.

3.3 Methodology for Context-Aware Services

Context-aware services incorporate models with the ability to analyze existing dependencies, predict causes and effects and provide decision making capabilities (which information services are relevant, how often are they needed, etc). For this purpose, we have designed a context-aware model which is a combination of syntactic, semantic and pragmatic services. The context-aware model supports users of hydrologic process model simulations in setting up inputs and interpreting results, in dynamically forecasting water quality problems and in alerting users.

The actual process of combining data and services over the Web is difficult and often requires specialized human expert knowledge. We use a knowledge-based approach to this problem. We collect, represent and classify data and service composition knowledge according to syntactic criteria (*How many inputs? What are the formats?*), semantic criteria (*Can this service function as part of the required service flow?*) and pragmatic criteria (*Does this service do what is needed by the user? Where is the user located?*). Especially the third category of *pragmatic knowledge* is a focus of our research. In our approach, pragmatic knowledge includes

¹ <ftp://ftp.wcc.nrcs.usda.gov/data/climate/basinreports/idaho/wy2004/baprid11.txt>.

context knowledge (*In what situation should this service be used?*). We now discuss the use of these three kinds of knowledge from the point of view of geospatial data and service integration.

Syntactic Model: Context-aware geospatial services require effective and dynamic integration of geospatial information. Such services (e.g., geospatial search, visualization, prediction and validation) deal with diverse data such as topographic maps, satellite images, and data about surface and ground water hydrology, contaminations, etc. from various geospatial databases and the Web. Identifying data from distributed and diverse resources, transforming them to compatible formats and integrating them for use in specific contexts are important steps for providing context-aware geospatial services. The data heterogeneity is caused by the diversity of data formats: map data (DXF, TIFF, etc.), attribute data, geospatial data (e.g., HPS) classification schemes (e.g. hydrologic units and zoning), and metadata (e.g. FGDC).

We view a service in the same way as a component in a data flow. Thus it has input and output ports. These ports expect data of certain types (e.g., *text* vs. *dxg* format) and meanings (numbers representing *yards* vs. numbers representing *temperatures*). Two services can only be composed if the input ports of one of them can be fed by the output ports of other services (allowing for user input and output). Thus, the first condition for matching services is that syntactic constraints be met. This includes the numbers and types of the inputs and outputs that need to be matched.

Let us assume that we have chosen a service (*Map* service) at the beginning of the flow and are looking for the next, syntactically compatible, service (*Water Quality* service) to concatenate to it. Whenever the syntactic constraints on compositionality cannot be met (for example, there is no water data available for a selected region), then a service is not composable. As another example, when a hydrologic Web service characterizes water by two parameters (*pH value* and *temperature*) and another service expects as input these two parameters and a third one (*water hardness*) then these two services cannot be composed due to syntactic reasons.

Semantic Model: From the semantic integration perspective, for a user's query, we investigate what data sets are most relevant to the given contexts and hydrologic models. Many hydrologic model simulators have been developed (including climate models, land-surface models, surface water models and groundwater models). Some models are well known and generally accepted by many research groups, while other models are not well known but provide more comprehensive outputs for specific purposes.

Web service compositionality often depends on the proper order of two services. For instance, if a new well has to be drilled, it is first required to obtain the groundwater and population data of the area, and then to locate possible sites for the new well. If this order is not respected, the two services (data retrieval and geospatial search) cannot be composed. This ordering relationship therefore constrains the compositionality of services.

Multiple hydrologic models can be integrated into our system *HIS-KCWater*. In order to simulate water conditions with a

hydrologic model for the locations for which no adequate data are currently available, we need to check whether data can be found locally or retrieved from the Web. In case that a hydrologic model is chosen, a coupled land-surface and groundwater model or a groundwater model can be selected to estimate the required data for a new well. We search the Web for hydrologic datasets and transform them into hydrologic model input data. We mainly use water data provided by the Department of Natural Resources in Missouri. Following our previous work on service flows [5], our techniques were applied to real data. This task confirms that our techniques can be used for groundwater studies and are compatible with current hydrologic datasets on the Web and with various hydrologic models.

The semantic composition rules often require extensive domain expert knowledge, such as knowledge of government regulations and policies, environmental statistics and processes, etc. For example, one important parameter to determine the quality of water is nutrient concentration. The nutrient concentration is often strongly related to other data such as pH value, temperature, discharge, turbidity, etc., so that a predication simulation can be done using data stored in the USGS database.

Pragmatic Model: Pragmatic knowledge is about the situations (context) in which a service should be used. Context-awareness is an integral part of service composition. The concept of context has been used in many different fields. Dey [6] defined it as a piece of information that can be used to characterize the situation of a participant in an interaction. Similarly, [4] defined context as location, environment, and identity of people and time. By sensing context information, present users different information, or modify their services according to changes in the environment. As the situation of the service requester changes with time and location, the new situation should be reassessed and services should be selected based on the context. In our case, the context model is comprised of location and time embedded with the query and feedback features. As such, pragmatic integration is not about the service itself, but about the way the service relates to the satisfaction of the goals of the user. In a drought case, monitoring the drought is critical for identifying intensive drought areas, ranging from either drying out and possibly heading for a drought, recovering but not yet back to normal, and suffering long-term impacts such as low reservoir levels².

Consider a scenario that requires a pragmatic service composition. The queries are: *Could you identify some effects of the recent drought in central Missouri? What do you predict will be the potential changes in groundwater levels? Are there any impacts on growing subsistence crops? How does it affect the population and water quality in the area?* The consequences and impacts of a drought may be analyzed considering water quality features (e.g., water temperature, contamination), agricultural features (e.g., crops or grasslands) and water supplies (e.g., rivers, or reservoirs). In the first place, the information service flow should identify the

² <http://www.drought.unl.edu/dm/archive/99/classify.htm>

specific drought in terms of location and time, possibly by invoking the U.S. Drought Monitor service². Then the flow may use this spatial and temporal data set as an input to query services for the fields that are known to be impacted by droughts.

The difficulties that pragmatic integration must overcome in linking these services together lie in the various levels of aggregation of temporal and geographical data, in incompatible geographical division systems (e.g. climatic, geological), in the possibility that data may not be available for all times and all areas, in the necessity to extract some information from maps, and in the need for reference values (e.g. averages). Additionally, the same concepts can be given different definitions by different services. For instance the USGS defines a drought in terms of percentiles below normal 7-day stream flow, while the U.S. Drought Monitor uses the more sophisticated Palmer Drought Index. In our scenario, additional information would associate the water supplies and the drinking water systems, in order to identify the populations that are exposed to drinking water shortages. Additional indicators are often needed in the West, where snowfall has a strong bearing on water supplies.

This example showed the need for pragmatic knowledge for dynamic service composition dealing with multiple domains. The scope of this analysis could range from local or regional to national. For instance, there could be water shortages or crop failures within an area not designated as a drought zone, just as there could be locations with adequate water supplies in an area designated as drought zone. Also, depending on the contexts of queries about specific locations like southern Mississippi or central Missouri, or on specific situations such as flooding or drought, the information provided by such services might be significantly different. Similarly, some websites provide regional level geospatial information while others provide national level services. Pragmatic knowledge is therefore used to intimately connect a query to the context in which a service flow is executed.

3.4 Development of Context-aware Rules

There have been several successful expert system shells for representing rules in a knowledge base, e.g., OPS5, JESS and CLIPS. Currently, the Semantic Web has adopted RuleML³ as "unofficial" rule representation language because it is based on XML. However, SWRL⁴ is under development to provide an RDF – based Rule/Axiom/Logic layer on top of the current OWL based ontology layer of the Semantic Web. A knowledge representation formalism consists of explicit syntactic notations which allow simple inferences. For example, *Lake* `<rdfs:subClassOf>` *Surface Water* states that all features and properties of the class *Surface Water* should also be applied to the class *Lake*. However, in real-world applications we need more of such "inferential notations" to allow more powerful inferences to be drawn. For example, consider the following relationship:

Missouri `<emerg:suffersFrom>` *Drought*. Here, we want to attach an inference rule that if Missouri is suffering from a drought, then we also infer the fact that *Missouri* `<emerg:perform>` *GroundWaterSupplyTest*. This is to be interpreted as a request that a groundwater test be performed by routing a well reading order to a human operator. This kind of low-level domain specific inferencing requires additional explicit rules. The idea is to have an expert system shell automatically infer additional related facts based on a minimal input of explicit facts. Jena⁵ provides a rule-based engine with capabilities for forward/backward chaining. Following is a simple Jena rule: If the height of the Missouri river is greater than 30 feet, then the hydrologists in the Missouri Flood Warning department will send a warning message, Danger:

[*FloodWarningRule*:

([http://a.com/ontology#Missouri river](http://a.com/ontology#Missouri%20river) *Missouri*
[http://a.com/ontology#height content](http://a.com/ontology#height%20content) ?val), *greaterThan*(?val, 30)
(<http://a.com/ontology#Hydrologist>
<http://a.com/ontology#sendMessage>
<http://framework.org/FloodWarning#Danger>)]

All three types of integration (syntactic, semantic, pragmatic) require cross-referencing, and sometimes inferencing from data with regard to a common ontology. We believe that the syntactic and semantic integration of services can be performed more effectively due to the existence of the pragmatic integration model. Additional mapping information would allow associating the water supplies and the drinking water systems, in order to identify the populations that are exposed to drinking water shortages.

4. HIS-KCWATER SYSTEM

As a proof of concept, we have developed a Web-based system, called *HIS-KCWater* (Hydrologic Information System for Kansas City metro WATERsheds). *HIS-KCWater* is an automated real-time analysis and forecasting system for water quality in the watersheds covering the Kansas City metro area. The basic features of *HIS-KCWater* are 1) Real-time data analysis: Various water quality data such as pH, temperature, turbidity, and discharge, and surface and groundwater monitoring data are statistically and geospatially analyzed in a real-time manner, 2) Interactive MAP server: A Web-based GIS map server allows a user to easily access monitoring stations on the map (through Google map services) and obtain real-time data with analysis results in an interactive manner, 3) Context-aware Water Quality forecasting: Embedded regression modeling tools enable forecasting nutrient concentrations from real-time water quality data. This forecast informs users about a potential degradation of water quality, 4) Personalized service: Each user can set up an environment of data acquisition and analysis options for a specific watershed or monitoring station. All functions are user friendly so that a user can obtain the information with minimal knowledge of modeling and GIS.

³ Rule Markup Language. <http://www.dfki.uni-kl.de/ruleml/>.

⁴ SWRL: A Semantic Web Rule Language Combining OWL and RuleML. <http://www.daml.org/2003/11/swrl/>.

⁵ Jena 2 - A Semantic Web Framework . <http://www.hpl.hp.com/semweb/jena.htm>.

The **HIS-KCWater Domain** is composed of seven major watersheds over six metro counties in Missouri, Kansas and Iowa. The Missouri River and Kansas River flow through the Lower Missouri Crooked watershed and the Lower Kansas watershed. As a 'test bed' platform, the current version of HIS-KCWater provides various analysis tools for water quality in these two major watersheds using GIS maps and real-time water quality data provided by USGS and EPA. There are major boundaries between watersheds: Lower Missouri Crooked and Lower Kansas in HIS-KCWater. The current version of HIS-KCWater supports five stations: two in Kansas and three in Missouri in the Kansas City metro area. The real-time water quality data are discharge, water temperature, pH, specific conductivity, dissolved oxygen, and turbidity, which are all available at five stations.

The **HIS-KCWater prototype** system has been implemented in Visual Studio.NET, which allows us to wrap existing Windows applications (i.e., WARMF and BASINS) as Web services and integrate them with the *HIS-KCWater* system in a Service-oriented architecture. The system consists of five core services; Interactive Map service, Time Series Analyst service, Google Map service, Water Quality Forecasting service and Personalized services. These services are logically connected but physically distributed; the Map Server is implemented using Java with various open source libraries (including GDAL, OGR, Proj4, Shapelib, LibJPEG, LibPng) on a Linux platform (Fedora core 5 Server) and configured with Apache Web server for a Web based interface. Other services have been implemented using ASP .Net and C# in the VS .NET framework and run on MS-IIS server. The location database is implemented in MS Access.

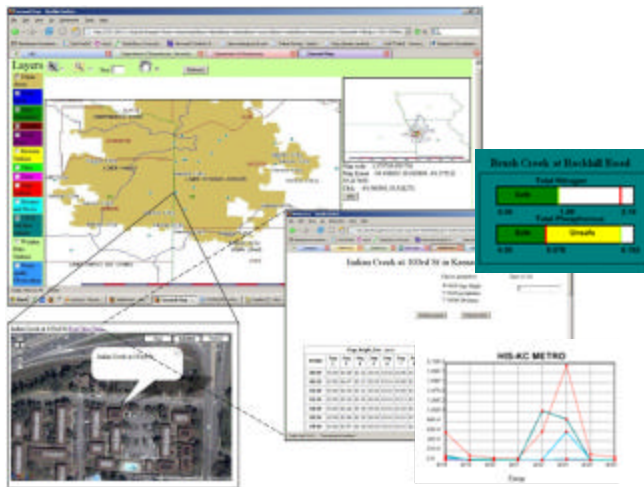


Figure 1. The HIS-Water System Interfaces

Figure 1 shows the screenshots of the interactive MAP server service for the real-time monitoring station at Indian Creek at 103rd Street in Kansas City, Missouri. If a user is interested in the water quality conditions at the specific area, she can easily search for its geographical location using integrated Google mapping and find real-time conditions of water quality such as pH, temperature, etc. In addition, the user can also estimate the expected concentration

of nutrients including nitrogen and phosphorous and predict whether the area is exposed to a potential threat of contamination from such NPS pollution. The user can compare individual water quality parameters at multiple stations in a given time range and obtain the spatial distribution of the parameters using the spatial analysis module. Context-aware water quality forecasts can be sent to registered users by email.

The **Interactive Map Server** has been developed using open sources, which allow to render spatial data on a multi-layered map and to provide interactive interfaces for selecting a specific region. The Map Server allows a user to zoom and pan through the layers of the map that she selected. The layers of spatial data in the Map Server include HUC watersheds, USGS real time stations, Weather data stations, Water Quality observation stations, etc. The spatial data for the various layers is represented in the ESRI format, using the Shapelib and MapServer script.

If the user clicks on a station for which we maintain real time data then she will get linked to the Time Series Analyst processing the data. If a user clicks any part of a map, a satellite view of that location can be provided by the Google Map Service. The Map Server has been incorporated with the Time Series Analyst and Google Map application. Besides watershed real time data analysis it provides additional geographical information, such as road, water station and residential area information.

Time Series Analyst provides users with analytical outcomes in graphical and table form. The user selects the station, parameters and time period and then the real time data will be fetched from the USGS Web server or local databases and displayed in the form of a table or graph. Real time information about various stations and various parameters can be analyzed and compared by plotting the graphs for values of requested parameters for the last seven days. For this simulation, the data have been extracted every 15 minutes from USGS for five selected sites.

In the *HIS-KCWater* system, several simulation models have been implemented to get real time measurements of Nitrogen, Phosphorous and Fecal caliform bacteria contents in the water. The concentration of nutrients in the water, such as nitrogen and phosphorous, is estimated by using linear regression models based on water quality parameters, including pH, temperature, discharge, turbidity, etc. In this approach, linear regression equations require measured water quality data as inputs to estimate the concentration of nutrients as an output. For example, we used a regression equation to estimate total nitrogen in water.

Context-aware Alert Service: An *HIS-KCWater* member can request the Alert Service by specifying particular types of data or specific stations or regions, ranges of selected data, and a mode of alert, such as email or phone call. The user can register the alerts such that if the value of a particular parameter is above or below a threshold value, the system will send the user an email. As mentioned previously, drinking water quality and safety regulations vary widely. In order to provide the context-aware Water Quality service, additional information is required such as

whether the user is in Missouri or Kansas, what the season is, whether the water is for drinking or other activities, where the activities are performed, etc.

HIS-KCWater is designed to intelligently forecast the potential degradation of water quality through the forecasting module. As seen in Table 1, each state has different criteria for water quality parameters. The pH and temperature of drinking water is important in determining water quality. A pH range of 6.0 to 9.0 is considered safe for drinking water in Missouri but Kansas is more stringent on the pH value (6.5 – 8.5). If any water quality parameter is in the range of ‘unsafe’ (Table 1), users can receive context-aware alerts and attempt to take corrective actions.

Table 1. Water Quality Regulation for Kansas and Missouri

Parameter	Unit	Missouri	Kansas
pH	pH	6 ~ 9	6.5 ~ 8.5
Turbidity	NTU	N/A	9.9
Dissolved Oxygen	mg/L	N/A	> 5.0
Temperature water (WT) natural air (CT)	Fahrenheit	WT<= 90 & WT- CT<5	WT<=90& WT- CT<3
Nitrogen	Stream & river	mg/L	< 2.18
	lake & reservoir		< 0.78
Phosphorous	Stream & river	mg/L	< 0.076
	Lake & reservoir		< 0.037
E.Coli Bacteria	col/100mL	126	N/A
Fecal Coliform Bacteria Recreation for 4/1-10/31	col/100mL	200	200
Fecal Coliform Bacteria Recreation for 11/1-3/31		1800	2000

Personalized Services: The HIS-KCWater system provides a number of Personalized Services for user friendly water management and forecasting of Kansas City watershed data. Member log in and profile services are provided for secure and customized access. The HIS-KCWater members can publish their data in the databases and integrate them with other data from different databases. They can request the data broadcasting service by specifying a particular type of data, specific stations or regions of their interest and also preferred hydrological simulation models.

5. Conclusion

In this paper, we presented a context-aware geospatial data and service integration framework for context aware services that utilize the analysis of data dependencies, identify relevant services and perform the services in order. In the HIS-KCWater system, we showed that 1) the integration of a water quality model and a real-time monitoring system enables one to predict a potential degradation of water quality, 2) the water quality analysis is context-aware and automated so that users do not need GIS knowledge to acquire real-time monitoring data, 3) the interpretation of the advanced analytical results about drinking water is likely to enhance the access to clean water.

6. REFERENCES

- [1] C. Aggarwal, F. Al-Garawi and P. Yu. (2001) Intelligent Crawling on the World Wide Web with Arbitrary Predicates", In Proceedings of the 10th International World Wide Web.
- [2] G. Adomavicius, R. Sankaranarayanan, S. Sen, A. Tuzhilin, (2005) Incorporating contextual information in recommender systems using a multidimensional approach, ACM Transactions on Information Systems, 23(1), 103 – 145.
- [3] T. Berners-Lee, J. Hendler, O. Lassila. (2001) The Semantic Web. Scientific American, 284(5):34-43.
- [4] P. J. Brown, J. D. Bovey, and X. Chen. (1997) Context-aware applications: From the laboratory to the marketplace. IEEE Personal Communications, 4(5):58-64.
- [5] S. Chun, Y. Lee, and J. Geller. (2004) Ontological and pragmatic knowledge management for web service composition. In 9th International Conference on Database Systems for Advanced Applications.
- [6] K. Dey, G. D. Abowd. (1999) Towards a better understanding of context and context awareness.
- [7] M. Donnelly and T. Bittner. (2005). Spatial relations between classes of individuals. In D. Mark and T. Cohn (eds.): Spatial Information Theory. Cognitive and Computational Foundations of Geographic Information Science. International Conference (COSIT 2005), 182 - 199.
- [8] D. Gruhl, L. Chavet, D. Gibson, J. Meyer, P. Pattanayak, A. Tomkins, J. Zien, J. (2004). How to build a WebFountain: An architecture for very large-scale text analytics. IBM Systems Journal - Utility Computing, 43(1).
- [9] S. Handschuh, S. Staab, A. Maedche, A. (2001) CREAM - creating relational metadata with a component-based. In Proceedings of the First International Conference on Knowledge Capture.
- [10] C. Jenkins, M. Jackson, P. Burden, J. Wallis (1999). Automatic RDF metadata generation for resource discovery. Proc. of the eighth international conference on World Wide Web. Toronto, Canada, pp. 1305-1320.
- [11] C.B. Jones, R. Purves, A. Ruas, M. Sanderson, M. Sester, M. van Kreveld, R. Weibel (2002) Spatial information retrieval and geographical ontologies. An overview of the SPIRIT project. SIGIR'02.
- [12] J.M. Pierre (2001) "On the Automated Classification of Web Sites", Linkoping Electronic Articles in Computer and Information Science, Vol. 6.
- [13] B. Smith and D. Mark (1998) "Ontology and Geographic Kinds," presented at International Symposium on Spatial Data Handling, Vancouver, BC, Canada.
- [14] H. Stuckenschmidt: Toward Multi-viewpoint Reasoning with OWL Ontologies (2006) ESWC Conference, pp. 259-272.
- [15] T. Tezuka, T. Kurashima, K. Tanaka (2006) Toward tighter integration of web search with a geographic information system. In Proceedings of the 15th International Conference on World Wide Web, 277-286.