'Photogrammetric Week 05' Dieter Fritsch, Ed. Wichmann Verlag, Heidelberg 2005.

OGC Specifications for Access to Distributed Geospatial Data

MATTHÄUS SCHILCHER, ANDREAS DONAUBAUER, München

ABSTRACT

One of the major international challenges of building Spatial Data Infrastructures (SDIs) is linking distributed, heterogeneous spatial information resources from different data providers in an application-oriented and user-oriented way. Exemplified by a utilities network information system for local governments this paper explains and discusses a way of overcoming these challenges based upon the specifications of the Open Geospatial Consortium (OGC). The utilities network information system is based on a unique multi-vendor OGC test platform. This platform has been set up in order to do research on interoperability and the usage of distributed spatial data and in order to assess and to promote the application of OGC specifications in practice by the Runder Tisch GIS e.V. ("Round Table Geographic Information Systems"), a vendor-neutral not-for-profit organization situated at the Munich University of Technology (Technische Universität München), Germany. This platform contains commercially available products from all major GIS vendors, products from smaller GIS vendors as well as Open Source Software implementing OGC Web Service specifications. Based on the experience gained in setting up the OGC test platform conclusions are drawn and recommendations for building SDIs are made regarding the benefits, the productivity, the applicability as well as the potentials and limitations of the OGC Web Services approach.

1. INTRODUCTION

1.1. The Problem: accessing distributed, heterogeneous spatial data

Many data providers in the private as well as in the public and the regulated sector have reached full coverage for their spatial databases. Consequently, today an increasing number of full-coverage, distributed, heterogeneous spatial databases exist as well as a vast amount of remote sensing data. The effort and the funds for creating and maintaining these datasets are often high but the data is still not being used to its full capacity. At the same time the traditional market segments for spatial information are saturated. In order to open up the market for new user groups, especially for non-GIS-experts, new innovative applications of existing spatial data sources are required [SCHILCHER 2004a]. An example for this kind of applications is the distributed utilities networks information system described in this paper. Applications of this kind typically rely on the combination of existing distributed spatial databases in order to derive information that meets the needs of a specific group of users. Aiming at reaching new user groups, it is critical to keep the initial technical and financial hurdles for using the technology as little as possible on both the potential user's and the provider's side. When considering that current GI systems require significant application knowledge and training as well as additional manual effort in integrating data from distributed sources it becomes clear that spatial data integration, the traditional approach to combining distributed, heterogeneous spatial databases, is only partly suited for reaching new user groups because it is demanding both technically and financially.

1.2. The solution: OGC Web Services

A new and promising approach to loosely or even spontaneously coupling independent, heterogeneous systems is introduced by the **web services technology** – a current "hot topic" in core information science. This technology allows for establishing interoperability between systems from different vendors. In the GIS domain, the Open Geospatial Consortium (OGC) is developing specifications for standardizing the interfaces of spatial web services. Web services implementing these specifications are called **OGC Web Services**.

The utilities network information system described in this paper makes use of this technology. The spatial information resources linked in this Spatial Data Infrastructure (SDI) are not combined by transferring and integrating spatial data from several source databases into one target database but by requesting information from several spatial web services via standardized software interfaces. Thus innovative internet applications can be developed as the mostly costly processes of data integration and data replication in the target database are no longer required for such applications.
Exemplified by the utilities networks information system mentioned above**,** the potentials and the benefits of OGC Web Services will be described.

1.3. Experience of the Technische Universität München regarding OGC Web Services

Since the year 2000 the Technische Universität München (TUM) and the Runder Tisch GIS e.V. ("Round Table Geographic Information Systems"), a vendor-neutral not-for-profit organization situated at the Technische Universität München, has been gaining experience in the field of applying, developing and coupling OGC Web Services from different vendors.
Research is focused on the following topics:

- simplified access and more efficient usage of existing spatial data,
- a new user profile (non-GIS-experts) for using distributed spatial data,
- chaining spatial web services in order to create services fulfilling larger tasks,
- potentials and limitations of the web services approach,
- multi-vendor interoperability by means of OGC specifications,
- profitability and business models for using distributed spatial data.

One of the projects carried out in the field of OGC web services is the "multi-vendor OGC test platform", containing GIS products from all leading GIS vendors as well as open source software that implement OGC specifications.

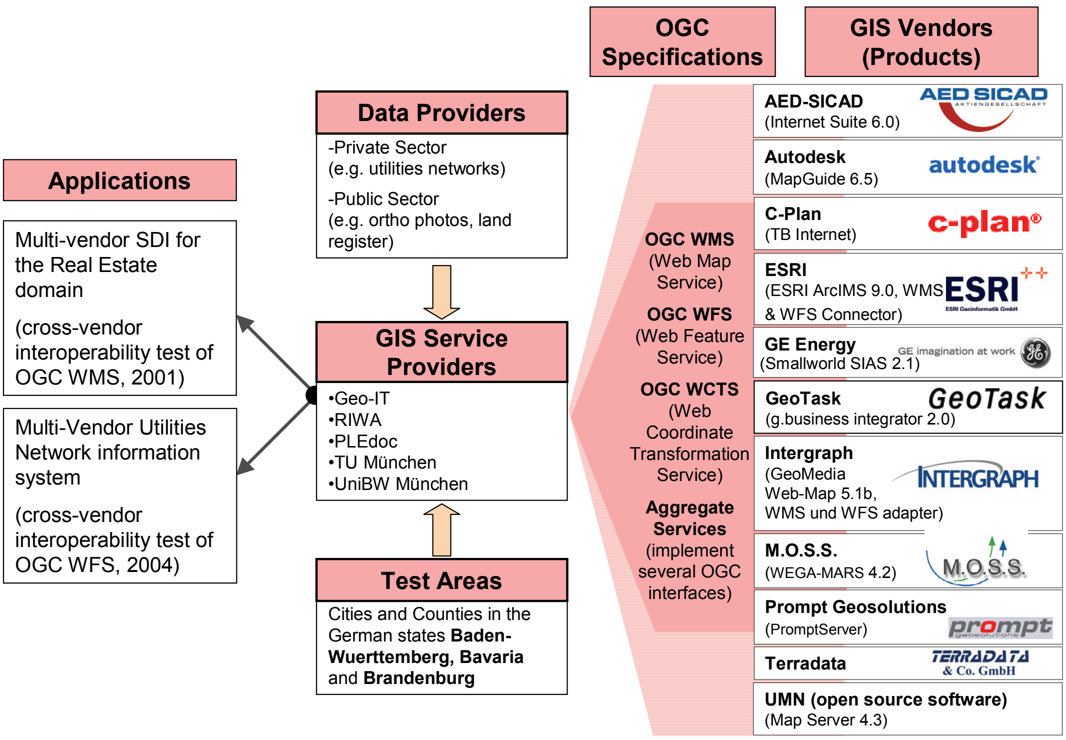

Figure 1: OGC Test Platform of the Runder Tisch GIS e.V.

With this platform, the Runder Tisch GIS e.V. aims to prove that:

- GIS from different vendors are able to interoperate based on OGC interfaces,
- users benefit from accessing distributed data provided by the private, public and regulated sector,
- the web services approach is capable of fulfilling the requirements of real-world scenarios and existing distributed, heterogeneous spatial data sources,
- the approach is applicable to the situation in Germany as well as in a cross-border environment.

Further characteristics of the OGC test platform are:

- vendor-neutrality,
- spanning various classes of business,
- cross-border scenarios,
- cooperation between data providers, GIS vendors, universities and the GIS service industry,
- research based on real-world applications that are of practical use to the GIS community,
- contributing to building SDIs,
- playing a leading role in the practice-oriented research on service chaining.

In 2003 the platform proved that GIS software from different vendors implementing the OGC Web Map Service Specification (WMS)[1] is able to interoperate and that the functionality provided by WMS is sufficient to fulfilling the requirements of a real world test case which was an application in the real estate business [SCHILCHER 2004b]. In contrast to earlier interoperability tests carried out by Runder Tisch GIS e.V., that test showed that the WMS specification is now stable and widely adopted by GIS vendors.
In 2004, the utilities networks information system described here was developed, proving multi-vendor interoperability based on the OGC Web Feature Service Implementation Specification (WFS)[2] [DONAUBAUER 2004b].

2. EXAMPLE OF USE: A MULTI-VENDOR UTILITIES NETWORKS INFORMATION SYSTEM BASED ON OGC WEB SERVICES

2.1. Scenario and Requirements

Operators of utility networks in Germany such as utilities companies and municipalities are legally obligated to answer requests of third parties (citizens, municipalities, construction and engineering companies as well as other utilities companies etc.) regarding the geographic position of their networks. For example whenever a municipality wants to dig up a road, all utilities companies in the relevant area must be asked if their networks will be affected by the planned roadworks. Subject to their size utilities companies must handle up to several thousand requests of such kind per year, nowadays usually coming in by fax or telephone.
Due to the multitude of networks (gas, electricity, sewerage, water supply, long-distance heating and communications networks) and the large number of prospective customers, offering a **one-stop**

[1] A Web Map Service (WMS) produces maps of georeferenced data. We define a "map" as a visual representation of spatial data; a map is not the data itself [OGC 2002a].

[2] In contrast to a Web Map Service (WMS), a Web Feature Service (WFS) [OGC 2002b] allows a client to request spatial data – not a map – from a service. The spatial data is encoded in the Geography Markup Language GML, another OGC specification [OGC 2002c].

service for obtaining information about all the utilities networks in a given area by just one request (Figure 2) seems to be a **promising business idea**.

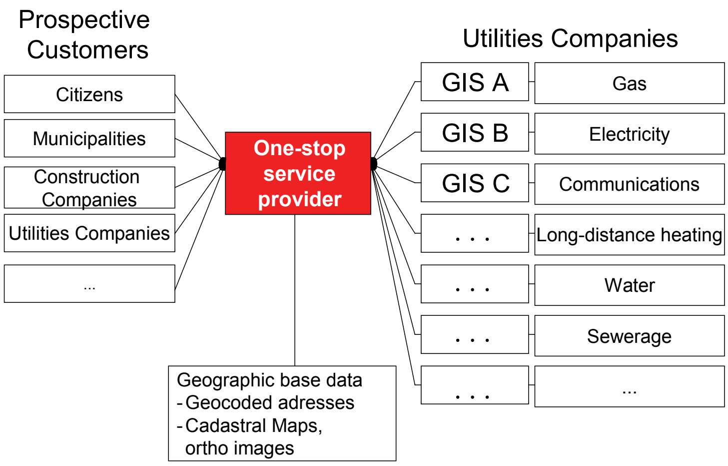

Figure 2: One-stop internet service for obtaining information about utilities networks across different utilities companies

Any organization that wants to set up such a service ("one-stop service provider" in Figure 2) must have access to the spatial data of the utilities companies as well as to geographic base data provided by other organizations such as national mapping agencies. When setting up the service the one-stop service provider is faced with two main difficulties: First, the data is **distributed** across several organizations and **heterogeneous** as it is stored in proprietary systems of diverse GIS vendors (symbolized by "GIS A", "GIS B" etc. in Figure 2). Second, utilities companies often **do not want to give their data to third party organizations**.

2.2. Approaches to combining data from distributed, heterogeneous sources

Basically there are two approaches to solving the problems of the distribution and heterogeneity of spatial data:

- **Approach A (classic): Data integration** (the service provider sets up a database that contains copies of the utilities companies' data and the geographic base data),
- **Approach B (new): Interoperability by means OGC Web Services** (the data remain in the systems of the utilities companies. These systems just give answers to requests that are issued via standardized web interfaces).

From the one-stop service provider's point of view **approach A is problem-prone** because of the following reasons:

- **Approach A** requires **utilities companies** to transfer their data to the service provider which they often dislike for competitive or legal reasons.
- **Data integration** is **time-consuming** and **expensive** and may cause information loss.
- The **service provider is in charge of keeping the integrated database up to date** which on the one hand means a **large effort** and on the other hand may lead to **problems regarding the liability of information** given to the customer because of out-dated data.

Based on these facts the Runder Tisch GIS decided in 2004 to implement an application based on approach B – accessing distributed GIS by means of OGC Web Services in order to demonstrate the feasibility of this approach.

2.3. System architecture and sample workflow

Figure 3 depicts the system architecture of the distributed utilities networks information system pointing out the OGC interfaces WMS and WFS.

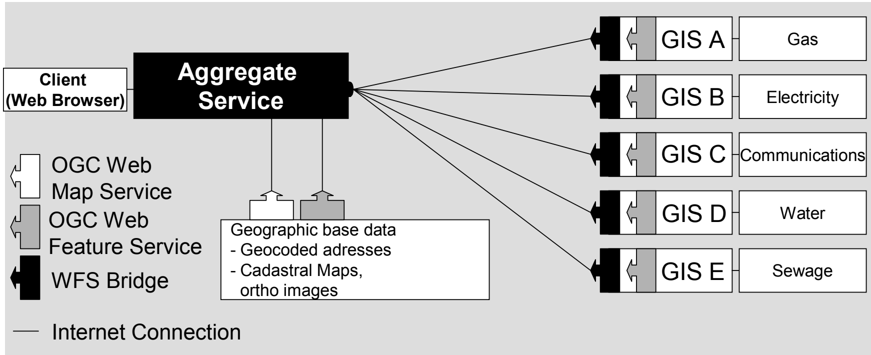

Figure 3: System architecture

Whereas the OGC Web Services (WMS and WFS) are used to access the distributed GI systems of the utilities companies and the producers of geographic base data, the application specific aggregate service serves for chaining the OGC Web Services so that the requirements of the users are met. Therefore an aggregate service has to fulfill the following tasks:

- It provides the user interface (e.g. a HTML document), as it accepts user input and transforms it into requests to OGC Web Services (OWS) and it presents (intermediary) results to the user for example.
- It transforms the output of one OWS into the input of the next service to be requested in the service chain.
- It combines the output of several OWS.
- Subject to the output of a OWS and rules defined by the developer of the aggregate service it branches the workflow.
- It reacts to malfunctions of OWS in the chain and to inconsistencies in the output of OWS.
- It hides the complexity of the overall system from the user, so the user has no awareness of the individual services answering his/her geospatial question.

The following figures represent a sample workflow for asking the system which utilities networks are affected by roadworks that are planned at an address specified by the user.

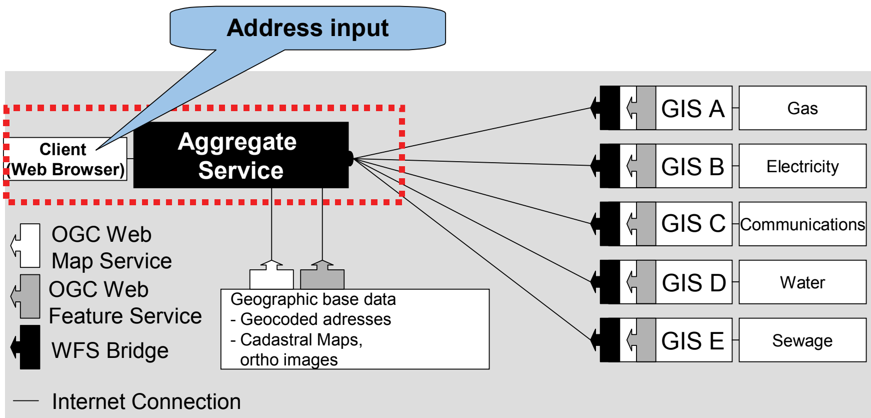

Figure 4: address input

The user does not have to install any software on his/her PC, all he/she needs is a web browser and access to the internet.
By entering an address (Figure 4) the user specifies the geographic location of the roadworks. The input mask for entering the street, house number and zip code is provided by the aggregate service.

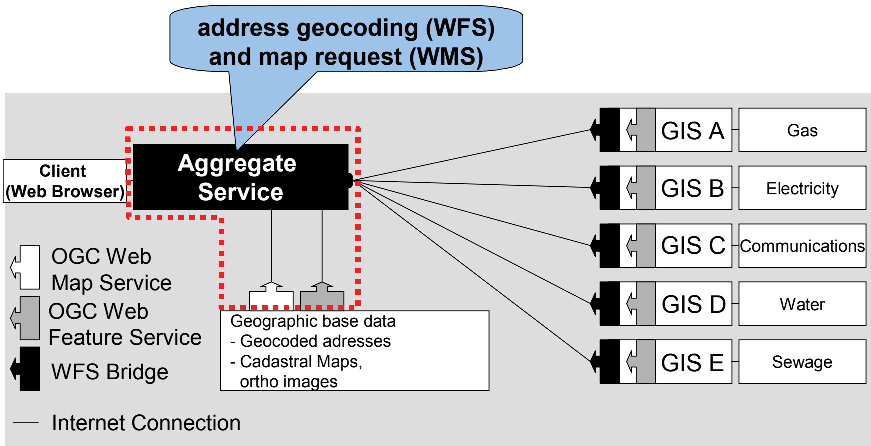

Figure 5: geocoding

The aggregate service transforms the address specified by the user into a request to a geocoding service (here an OGC Web Feature Service which encapsulates a database containing geocoded addresses). The geocoding service transforms the address into coordinates (Figure 5).

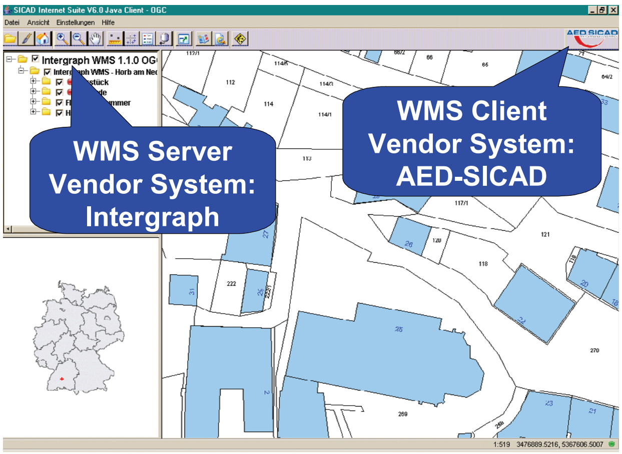

Figure 6: geographic base map presented in the user's web browser

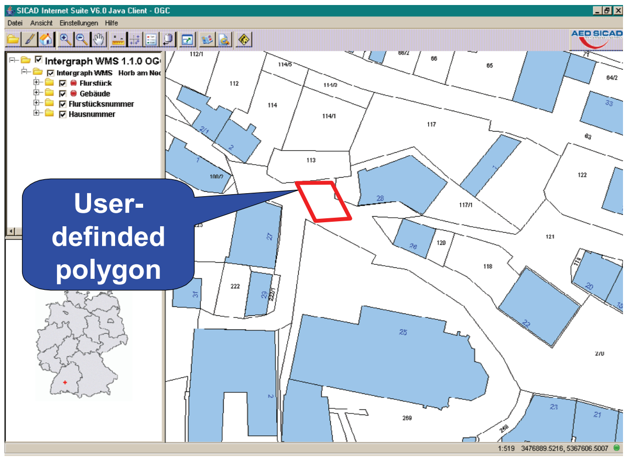

Figure 7: highlighting the location of the roadworks by drawing a polygon

Using these coordinates the aggregate service requests a base map from a WMS service and sends this map to the user's web browser (Figure 6).

The user then highlights the location of the roadworks by drawing a polygon on the base map (Figure 7). The polygon is then sent to the aggregate service by the user's web browser.

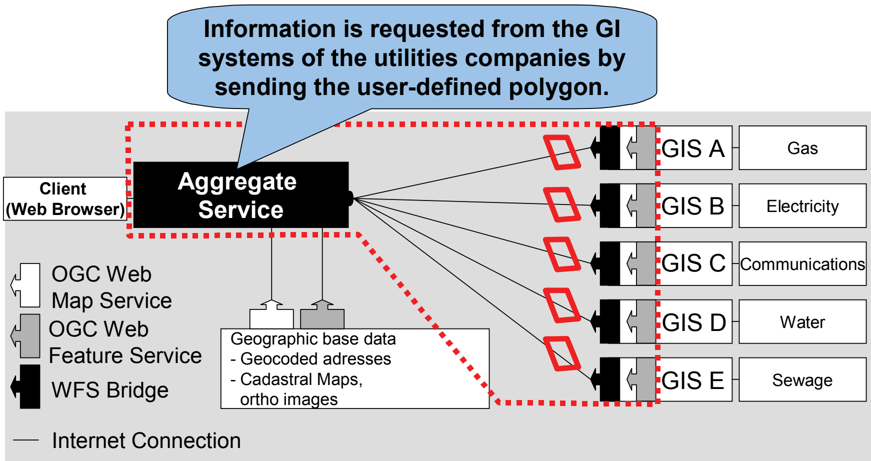

Figure 8: Requesting information from the utilities companies' GI systems from the vendors C-Plan, ESRI, GE Energy, Intergraph, M.O.S.S. (GIS vendors) and University of Minnesota (open source)

The aggregate service receives the polygon and sends it to the distributed GI systems of the utilities companies (WFS operation „GetFeature", spatial filter „Intersects" with the polygon being the argument of the filter). Between the aggregate service and the WFS interfaces of the utilities companies' GI systems there is a component called "WFS bridge" that acts as a WFS client and transforms the WFS responses (GML feature collections) into answers of the kind "utilities network is affected / not affected" and sends these answers to the aggregate service (Figure 8). The aggregate service assembles the answers of the utilities companies' distributed systems and sends the overall result to the user. Browsing the web page generated by the aggregate service, the user can then tell which utilities companies networks are affected by the roadworks he/she plans to carry out and which are not. The user retrieves this information with minimum effort because all the utilities companies are queried in one go.

3. BENEFITS OF THE OGC WEB SERVICES APPROACH

The above described example of use demonstrates the feasibility and the benefits of the OGC Web Services approach for combining distributed, heterogeneous GI systems. The main benefits are:

- **Minimized effort for using and combining spatial data on the users' side:** The end user neither needs to store and process spatial data nor does he/she have to be a GIS expert. The end user only has to have internet access and a web browser.
- **Reduced effort, lower costs and more efficiency on the service providers' side:** The service provider saves time and costs because the data integration work steps data collection, data transfer, format conversion, data preparation and data administration are no longer necessary in order to retrieve information from distributed data sources.
- **Increased information liability because of updateness:** Whenever a user requests information from the service provider, the original data sources of the data providers (utilities companies and providers of spatial base data in our example) are queried. This increases the quality of a service result because the risk of dealing with out-dated data is minimized.
- **Reusability of services:** As Geo Web Services can be combined easily, one and the same Geo Web Service can be reused in different contexts and applications once it has been published within an SDI.

4. LIMITATIONS OF THE OGC WEB SERVICES APPROACH

The research done by Runder Tisch GIS e.V. not only demonstrates the potentials of the OGC Web Services approach but also its limitations described in the following sections.

4.1. Practicability of current OGC Web Service Specifications

Some of the requirements defined by real-world applications cannot be met by the existing specifications for OGC Web Services. Known limitations concerning the practicability are:

- **Effort in combining different types of OGC Web Services:** Chaining different types of OGC Web Services (e.g. Gazetteer Service and Web Map Service) requires consistency between OGC specifications. At present an application developer who wants to chain different types of OGC Web Services has to deal with complications that arise from inconsistencies. In our example the aggregate service compensates for syntactic inconsistencies such as different representations of a Bounding Box in a WMS request and in a GML document (the response of a Web Feature Service / Gazetteer Service).
- **Effort in integrating OGC Web Services in a mainstream IT infrastructure:** As the current versions of OGC specifications do not support all the mainstream IT standards for web services (in particular the Simple Object Access Protocol SOAP), an integration in mainstream IT might be costly.
- **Limitations of distribution and modularity:** Tests carried out by the Technische Universität München [Donaubauer 2004a] showed that under certain conditions answering a geospatial question (e.g. "Which land parcels are adjacent to water bodies and are owned by municipality X?") based on distributed data storage and distributed processing is less efficient than retrieving the same information from an integrated database.

4.2. Acceptance and productivity of OGC web services

Figure 9 shows maturity and time to productivity of web services technology in general compared with other cutting-edge new technologies by means of the "Gartner's Hype Cycle", a diagram devised by the US research firm Gartner Group.

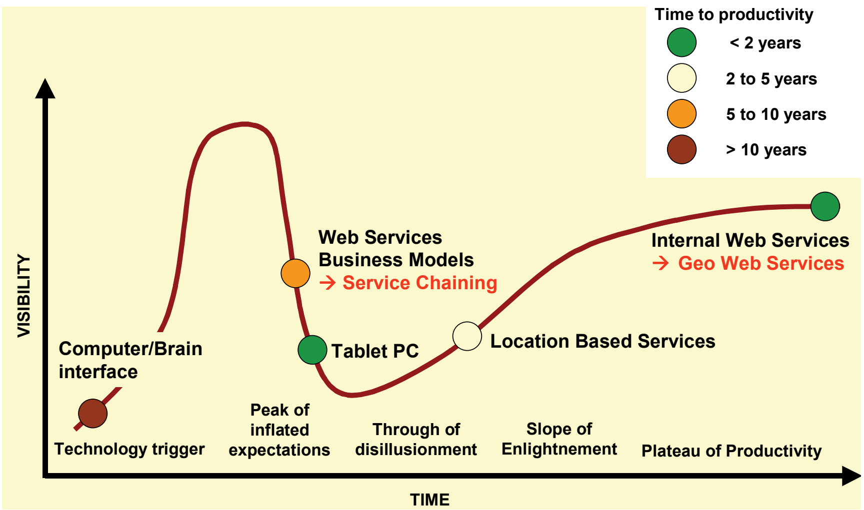

Figure 9: Gartner's Hype Cycle of Emerging Technologies
(source: Gartner Group 2004, modified)

For the GIS domain, the following statements can be made regarding the acceptance and productivity of standardized web services, namely the web services specifications of the OGC. The acceptance of OGC specifications in practice – by GIS vendors, providers of spatial data, service providers – is a prerequisite for the success of the approach described in this paper. The OGC web services approach will not be practically and economically advantageous until many GIS vendors provide their software with easy-to-use OGC interfaces on server and on client side and not until the major customers of the GIS vendors (e.g. national mapping agencies) demand and apply these interfaces. Some of the OGC Web Services specifications[3] are already mature and widely accepted by GIS vendors. Multi-vendor interoperability could especially be demonstrated for systems implementing the **Web Map Service Implementation Specification WMS**. It could also be proved that solutions based on the WMS specification meet the requirements of real-world use cases (SCHILCHER 2004). As far as the **Web Feature Service Implementation Specification WFS** is concerned, the OGC test platform described above shows that this standard is also but sometimes fragmentary implemented by the leading GIS vendors. Today, most of the OGC implementations **provide read-only access**. The possibility of a write access by means of OGC Web Services, provided by a so called transactional WFS, is rarely implemented by GIS vendors. Reasons for that are on the one hand the complexity of a write access to a spatial database via the internet (transaction management etc.) and on the other hand a lack of demand for this functionality from the operators of OGC web services. In general data providers are still apprehensive about allowing online access to their data via the internet. A reason for that can be found in the absence of security and access control (digital rights management) in the current OGC specifications. The OGC has just begun working on these subjects.

4.3. Limitations in functionality

Even though the implementation of more or less simple but profitable web applications based on OGC web services is not a problem any more, there is still a lack of analysis functionality in the current OGC specification. Metric queries (calculation the area, perimeter etc. of geometric properties of features) and analysis methods that derive new features from a given set of features such as polygon overlay are not available in approved OGC specifications[4].

The principle of the OGC web services approach is to hide the arbitrary complex internal structures of a system from the user while providing access to the system by the relatively simple structures of standardized interfaces. On the one hand this is an advantage because it allows for interoperability. On the other hand an OGC web service also hides the conceptual schema of the spatial data encapsulated by the service although this information would be of interest to a knowledgeable user who wants to process the data delivered by an OGC web service in his/her own GI system. In order to perform complex data analysis the structure of features, their relationships and their semantics must be known. This information can be derived from an exact and unambiguous description of the conceptual schema currently not provided by OGC web services[5].

---

[3] A list of all OGC specifications can be found on http://www.opengeospatial.org/specs/
The most important specifications for accessing remote sening data are the Web Map Service and the Web Coverage Service Implementation specification.

[4] The TU München submitted a draft specification for a so called "Web Spatial Analysis Service" to the OGC aiming to close this gap in the OGC specification.

[5] Research has shown that XML schema, the language that is used to describe the structure of features in the transfer format of OGC Web Services (GML) is not unambiguous (see SHI 2004).

5. CONCLUSIONS

Taking into account the current potentials and limitations of OGC web services the following preferred classes of applications can be identified:

- **Applications in a multi-vendor environment where data storage facilities and GIS knowledge on the users' side cannot be implied**. This is the case for almost all new user groups of existing spatial data.
- Applications where **data updateness is critical** (see the example of use described in this paper).
- Applications **in a multi-vendor environment** that require access to a **wide range of distributed data sources mixing spatial and non-spatial attributes (**this is true e.g. for applications in the field of urban planning or real estate assessment**).**
- Applications **in a multi-vendor environment** that rely on a **combination of remote and local data** (e.g. topographic base data provided by a national mapping agency combined with local data of a municipality).
- Applications such as the management of natural and man-made hazards that require **ad hoc access to a combination of distributed heterogeneous spatial data sources**.

Considering that all the items in the above list are characteristic for any Spatial Data Infrastructure (SDI), it becomes clear that **OGC web services are viable technology for developing SDIs**.

The OGC test platform described in this paper proves that the **technology and the standards for implementing SDIs are already available** and implemented in current GIS software.

However, the current situation in SDI development can be described as follows:

- a relatively **high level of attention** for the topic amongst GIS experts is observed,
- there are **high expectations** but
- **there is very little awareness amongst GIS users and GIS non-experts** (the main target user group!) and
- the **benefits** of SDI are **not proved yet,**
- **business models** are lacking so far [MICUS 2004].

Thus, more demonstration projects like the one described in this paper are needed in order to prove the benefits of SDI, to increase the awareness amongst GIS users and to leverage this fruitful technology.

6. REFERENCES

Donaubauer, A. (2004) [DONAUBAUER 2004a]: Nutzung verteilter Geodatenbanken mittels standardisierter Geo Web Services, dissertation, TU München. Internet: http://tumb1.biblio.tu-muenchen.de/publ/diss/bv/2004/donaubauer.pdf

Donaubauer, A.; Fischer, F.; Huber, A.; Müller, S.; Plabst, S., Straub, F. (2004) [DONAUBAUER 2004b]: Leitungsauskunft aus verteilten GIS. Project report of the Runder Tisch GIS e.V., München. Internet: http://www.rundertischgis.de

Fornefeld, M.; Oefinger, P.; Jaenicke, K. (2004) [MICUS 2004]: Nutzen von Geodateninfrastrukturen. MICUS Management Consulting GmbH (Hrsg.).
Internet: http://www.micus.de/54_gdistudie.html

Open Geospatial Consortium (2002) [OGC 2002a]: Web Map Service Implementation Specification, Version 1.1.1, OpenGIS project document 01-068r3.

Open Geospatial Consortium (2002) [OGC 2002b]: Web Feature Service Implementation Specification, Version 1.0.0, OpenGIS project document 02-058

Open Geospatial Consortium (2002) [OGC 2002c]: Geography Markup Language, Version 2.1.1, OpenGIS project document 02-099.

Schilcher, M., Aumann, G., Donaubauer, A., Matheus, A. (2004) [SCHILCHER 2004a]: Abschlussbericht zum High-Tech-Offensive-Projekt GeoPortal, Internet: http://www.rtg.bv.tum.de/index.php/article/archive/121

Schilcher, M., Teege, G., Donaubauer, A., Kunkel, Th. (2004) [SCHILCHER 2004b]: OGC Web Services zur interoperablen Nutzung verteilter Geodatenbanken für die Immobilienwirtschaft. In: Bernard, L., Fitzke, J. und R. Wagner (Hrsg.): Geodateninfrastrukturen. Grundlagen und Anwendungen. Heidelberg.

Shi, W. (2004) [SHI 2004]: Zum Modellbasierten Austausch von Geodaten auf Basis XML. Dissertation, Universität der Bundeswehr München.