INTEROPERABILITY IN GI SERVICE CHAINS – THE WAY FORWARD

Lars Bernard¹, Udo Einspanier¹, Michael Lutz¹, Clemens Portele²

¹Institute for Geoinformatics, University of Münster
Robert-Koch-Str. 26-28, D-48149 Münster, Germany
e-mail: {bernard, spanier, lutz}@ifgi.uni-muenster.de

²interactive instruments GmbH
Trierer Straße 70-72, D-53115 Bonn, Germany
e-mail: portele@interactive-instruments.de

1. INTRODUCTION

Currently GI-Science is dominated by a paradigm shift – from monolithic to open and interoperable GIS, and from standardized data formats towards specifications of GI service interfaces (Abel et al. 1999, Groot and McLaughin 2000). The number of GI services available on the web is continually increasing. Interoperability is supposed to be ensured by a number of specification efforts, in the geo domain most prominently by ISO/TC 211 and the OpenGIS Consortium (OGC). These efforts may ultimately allow the (semi-) automatic composition of several independent services into a complex service chain. The issue of service composition is most notably addressed in a standard describing an abstract service architecture (ISO/TC-211 and OGC 2002).

These standardization efforts notwithstanding, (semi-) automatic service composition is still very rare. One reason for this is that the standards described above – while resolving syntactic heterogeneities – still do not solve semantic heterogeneity (Bishr 1997, Visser et al. 2000).

In this paper we present a use case for which a complex service chain is composed of a number of services that have not been intended to be applied in this specific chain. Based on this service chain – as a state-of-the art example in GI service chaining – a number of restrictions will be discussed. A second scenario will show how to answer the use case’s question in a more dynamic and flexible way. The semantic problems arising in this scenario are the focus of the discussion. Finally possible solutions will be addressed and those problems for which a solution is currently missing will be identified.

2. THE USE CASE – ESTIMATING ROAD BLOCKAGE AFTER STORMS

The use case described here is set in the area of disaster management and mitigation. In December 1999 the winter storm “Lothar” caused an accumulation of about 150 million solid cubic metres in windfall timber in central Europe (Gemeindetag Baden-Württemberg 2000).

The main actor in the use case is the person in the authority responsible for ensuring road safety. In order to coordinate the clearing operations effectively a rough estimation about which roads are most likely to be affected by fallen timber is required. This estimation is based on the assumption that those forest stands with overaged trees be most strongly affected by storms. At what age a tree can be called overaged depends on its species. The conventional way to answer this question would be to use a GIS for performing selections on
a forestry layer and overlaying the result with road data. However, if the user does not have access to local data sources containing the necessary information (e.g. age of trees) the question cannot be answered.

In contrast, the user can apply web based GI services providing and displaying the required information. The following chapters describe two scenarios of how these GI services could be combined to form a complex service chain. First, the state-of-the-art scenario is described where generic Web Feature Services (WFS) and Web Map Services supporting Styled Layer Descriptor (WMS/SLD) are statically chained to form a complex service for estimating road blockage after storms. This scenario was realized in the just finished GDI NRW Testbed II, the second testbed of the Geodata Infrastructure Initiative of North Rhine-Westphalia (Bernard 2002). Starting from this scenario, a future scenario is developed describing a more flexible way of service chaining.

3. STATE OF THE ART – CHAINING DISTRIBUTED WEB FEATURE AND WEB MAP SERVICES

In the context of the GDI NRW Testbed II a "road blockage" service for the use case described above was implemented (http://xtra.interactive-instruments.de/demo/demo-wfs.html). It lets the user select a lower age limit for trees, a certain tree species on which to constrain the query and as an option a maximum number of forest features to be queried (Fig. 1). The service returns a map of the road network combined with a topographic map with those roads highlighted for which a blockage is most likely (Fig. 2).

![Fig. 1 Query criteria in the user interface of the "road blockage" service: (1) Minimum age of trees, (2) Tree species, (3) Optional maximum number of forest features](image)

According to (ISO/TC-211 and OGC 2002) the "road blockage" service represents an aggregate service (opaque chaining). In a black box manner it accesses a number of OGC-conformant WFS (OGC 2002b) and WMS/SLD (OGC 2002c, OGC 2002a) provided by different members of the GDI NRW:

- a WFS hosted by the Institute for Geoinformatics in Münster (http://lfgi.uni-muenster.de) serving forest features from a database provided by the North Rhine-Westphalian Department for Forestry (WFS-lFGI),
- a WFS hosted by interactive instruments (http://www.interactive-instruments.de/) serving road features from a database containing the North Rhine-Westphalian road network (WFS-II),
a WMS hosted by the North Rhine-Westphalian agency for data processing and statistics (LDS, http://www.lds.nrw.de/) providing topographic maps (WMS-LDS), and

- a WMS/SLD hosted by interactive instruments for displaying the above-mentioned features in a map (WMS-ii).

![Image of a map highlighting roads]

**Fig. 2** Result of "road blockage" service: A map highlighting those roads that are most likely to have been affected by the storm.

The flow of information is depicted in Fig. 3. The "road blockage" service contains all the application logic and is thus solely responsible for controlling the flow of information. In addition it is responsible for performing a number of tasks, inter alia,

- translating the user input into the specific classification scheme of the forestry data and generating the WFS query based on the translated user input (OGC 2001) (3),
- accessing the forestry WFS-IfGI using the generated query (4–5),
- extracting the forest features' geometries from the forestry WFS-IfGI response and using them for constructing the query against the road WFS-ii (6)
- providing the WMS-ii with the generated query against the road WFS-ii (via Styled Layer Descriptor, (OGC 2002a)) (7) and the queries for the road data map layer and the topographic map layer (8–9), and
- displaying the resulting map to the user.
As the services used in the chain were not developed specifically for the service chain described, this scenario illustrates the benefits of interoperability standards, at the same time representing the state of the art in the area of GI service chaining. However, the scenario also has a number of limitations.

In order to be able to implement the client application the provider had to have knowledge of the services’ existence, their capabilities and their application schemas (Fig. 4). The client application only works with these specific services and application schemas. Furthermore it cannot be reused for the execution of different chains as its workflow management is tied to this specific chain.

These limitations are addressed in a second, more flexible scenario, which is presented in the following section.
Fig. 4 GetFeature request against the forestry WFS-IFGI. The types and properties of the features served by the WFS-IFGI (marked in bold) have to be known to construct a filter for the selection of overaged tree stands.

4. THE WAY FORWARD – AD-HOC SERVICE CHAINING

The second scenario describes a workflow-managed (translucent) service chain (ISO/TC-211 and OGC 2002). While the ISO/TC 211 standard only states that in such a scenario a pre-defined chain is selected by the user, we further assume that the chain definition does not exist at the time the user poses his question, but that it is assembled in an ad-hoc fashion by a workflow (composition) service. Accordingly, the user should be able to specify his question in a generic user interface, provided by appropriate human interaction services. This interface could behave similar like for instance, query masks or analysis wizards in current GIS (Fig. 5, left), or - as a more futuristic version - also look like user interface of internet search engines (Fig. 5, right). Catalogue services should help the user to further specify his question by providing (semantically rich) information on available services. Workflow services will then be responsible for composing and executing a chain that answers the question thus, that the semantics of the answer match the semantics specified by the user. This includes the subtasks of translating the user’s question into formal requirements, generating a solution strategy to solve a complex problem with several smaller tasks, matching the requirements to the capabilities of the available services and offering some quality measurements to evaluate the fitness of use of the service chain’s results.

The flow of information is depicted in Fig. 6. First, the user has to enter his query. In order to assist the user in this task, this will have to be a highly interactive and iterative rather than a linear process (1a-d). In order to do that, capability descriptions of available services are matched against the user requirements. That will require information on the available services and data from the catalogues as well as formalized domain knowledge. The query entered by the user is then translated into a workflow describing a service chain (2). Queries against the components of the chain have also to be formulated in appropriate query languages (3).
The execution of the actual service chain is controlled by the workflow management service. After invoking the chain execution (4) the flow is the same as in the first scenario with the exception that the services notify the user of their actions (6b, 9b, 11b, 12b).

**Fig. 5** GIS-like (left, here Desktop-GIS ArcView™) or Google™-like (right) query interface for constructing GI-service chains

**Fig. 6** Flow of information using translucent service chaining
5. INTEROPERABILITY ISSUES AND POSSIBLE SOLUTIONS

While the scenario described in the previous section would greatly increase flexibility, a number of interoperability issues have to be addressed before it can be implemented. The problems fall into two broad categories, supporting the user in specifying his query and automatically matching (user or service) requirements against (data or service) capabilities. In order to solve them, several human interaction services will be required. Consequently, these services, having so far been treated as a black box, will have to be further specified, ultimately leading to a more detailed (white box) architecture for ad-hoc workflow-managed service chaining. In order to support the user in specifying his query the following questions have to be addressed:

- **Interaction between user, human interaction services and catalogues.** The workflow between the user, the human interaction services and the catalogues has to be defined. This includes the specification of the components necessary for this workflow (e.g. a user interface, a service for controlling the query workflow and a catalogue service). The question of the user interface design should also be addressed.

- **Translation of user input into formal requirements.** In order to apply automatic matchmaking algorithms in the search for appropriate services and data, the workflow description must also contain formal specifications of the services needed for answering the query. Finding formalism for these specifications and methods for translating the user input into them are further questions to be addressed. These questions are closely linked to the question of how to specify service capabilities and the methods employed in the matchmaking process (see below). As for describing web services requirements DAML-S (The DAML Services Coalition 2002) could be used for formally representing user queries (Paolucci et al. 2002). Visual query wizards currently available in GIS (Figure 5) provide a starting point for developing tools, which help to formalize the users’ queries.

- **Composition of service chain.** The query entered by the user subsequently has to be transformed into a service chain. Methods for partitioning the user's query to provide a sequence of simpler steps have to be developed. A formalism for representing the resulting workflow has also to be examined. The emerging industry standard BPEL4WS (Business Process Execution Language for Web Services) may serve as a first approach.

Once the requirements for a service to help in answering the user’s question have been formally specified, these have to be matched against the capabilities advertised in the catalogue. Services that are considered appropriate might require additional input, so that for composing the service chain, service requirements also have to be matched against (service or data) capabilities. More specifically the following issues have to be solved:

1. **Representation of static and dynamic aspects of geoinformation.** Functional equivalence between features is a key to enable interoperable open GI infrastructures (ISO/TC-211 2001b). It should be possible to express valid operations for particular geographic features in a specific state – as, for instance, discussed already by Kuhn (1995), Bernard and Krüger (2000) – and to find out if a particular service instance can operate on supplied geodata. Following ISO/TC 211 this information is modelled in an application schema: “An application schema defines: content and structure of data; specifications of operations for manipulating and processing data by an application” (ISO/TC-211 2001a). However, while the General Feature Model of ISO allows for modelling this information, current (XML-based) ISO/TC 211 and OGC standards for encoding application schemas provide only mere data transfer formats limited to static feature characteristics.
2. **Service and Data Catalogues.** Catalogues provide metadata libraries that allow to search for appropriate information. Currently there are two different standards for data and service metadata, leading to the separation of static and dynamic geoinformation aspects on the metadata level. Proprietary solutions have to be developed to integrate these two metadata schemas (one approach has been realized in GDI NRW Testbed II). Additionally, there is no query facility defined for the application schemas, inhibiting information retrieval based on constraints on feature type characteristics.

3. **Formal representation of service (and data) capabilities.** DAML-S (The DAML Services Coalition 2002) could be used as a formalism for representing service capabilities. However, the specification is still only available as a draft (currently version 0.7) with important sections still to be completed. Depending on the matchmaking methods used, it will have to be assessed for its fitness for use.

4. **Semantic Translation.** One important task in both matchmaking and workflow composition will be semantic translation between concepts from different domains. This should also include a mapping from the user requirements specification to a query for that can be sent to a specific service using its query language (Visser et al. 2000).

5. **Matchmaking Algorithms.** A number of algorithms for matching requirements to capabilities have been developed in the area of information agents (e.g. Sycara et al. (1999)) and recently been adapted for use with web services and UDDI catalogues using service descriptions in DAML-S (Paolucci et al. 2002). However, these approaches still seem to assume a relatively closed-world.

6. **Quality/degree of match.** The matchmaking algorithms described above also propose measures for the degree of match. In Sycara et al. (1999) a semantic distance between concepts, whose computation is based on a semantic network, is proposed. In contrast, Paolucci et al. (2002) suggest a discrete scale with four categories for measuring the degree of match. A combination or integration of these measures on GI quality measures, like spatial and temporal correctness, appropriate spatial and temporal scale, etc. is still to be done.

7. **Integrate matchmaking with user support.** Last but not least it should be investigated how matchmaking can be integrated with the process of query formulation (Wache et al. 2002). This will have strong impacts on how the workflow of the querying process communication will be organised.

6. **CONCLUSIONS AND FUTURE WORK**

Current GI standards developed by ISO/TC 211 and OGC define service interfaces for simple geoprocessing tasks like mapping, metadata and feature access. They are used as a basis to build distributed interoperable GI infrastructures accessible through the web. More complex tasks can be solved by chaining several of these simple services, tackling user problems outside the scope of one simple service that would otherwise require a trained GIS operator and local geodata. As an example, a "road blockage" service is presented that was realised in the context of GDI NRW testbed 2. However, GI service chain composition is still very limited.

This paper identifies a list of research and development issues, that need to be tackled to realize a workflow managed – semi-automatic or automatic – chaining of GI services. Our future activities focus on the study of current solutions to handle the first four issues of the list formulated in the previous section; thus, forming the basis for the design of appropriate match making algorithms. Further research will especially explore whether current approaches and concepts out of the general web service and semantic web domain are utilizable for GI service chains.
7. ACKNOWLEDGEMENTS
The work presented in this paper has been partially supported by the German Ministry for Education and
Science as part of the GEOTECHNOLOGIEN program and can be referenced as publication no. GEOTECH-14.
Furthermore parts of the work have been funded by the European Commission through the ACE-GIS project.

8. REFERENCES
Ruiz, Gould, M. & Ramon, J. (Eds.). 5th AGILE Conference on Geographic Information Science
Universität Wageningen.
Verheerende Stürme in Baden-Württemberg. Zeitschrift für die Städte und Gemeinden für
Stadträte, Gemeinderäte und Ortschaftsräte (Organ des Gemeindetags Baden-Württemberg).
2000.
Service Architecture Vs. 4.3. Draft Version. International Organization for Standardization &
OpenGIS Consortium.
(ISWC2002), Sardinia, Italy, Springer: 333-347.
0.7.
efficiently: Sharing data and knowledge from heterogeneous sources. In: C. Rautenstrauch &
Patig, S. (Eds.), Environmental Information Systems in Industry and Public Administration. Hershey
Florida Artificial Intelligence Research Society Conference (FLAIRS), Pensacola, FL, USA: 445 -
449.