A Classification Framework for Approaches to Achieving Semantic Interoperability between GI Web Services

Michael Lutz, Catharina Riedemann, and Florian Probst

Institute for Geoinformatics, University of Münster, Robert-Koch-Str. 26-28,
48149 Münster, Germany
{lutzm,riedemann,probst}@ifgi.uni-muenster.de

Abstract. The discovery of services that are appropriate for answering a given question is a crucial task in the open and distributed environment of web services for geographic information. In order to find these services the concepts underlying their implementation have to be matched against the requirements resulting from the question. It is in this matchmaking process where semantic heterogeneity has to be tackled. Whether semantic interoperability can be achieved depends on the quality of the information available to the matchmaker on the semantics of requirements and resources. The explicitness, structuring and formality of this information can differ considerably leading to different types of matchmaking. In this paper a framework is presented for classifying the approaches that are currently employed or proposed for achieving semantic interoperability according to these criteria. The application of the framework is illustrated by analyzing possible solutions to three examples of semantic interoperability problems.

1 Introduction

Geographic information science is currently characterized by a paradigm shift – from providing theories for monolithic systems to theories for open and distributed GIS and their use processes. With this comes a move from standardized data formats to specifications of geographic information (GI) service interfaces [1, 2]. In practice, the number of GI services available on the web is rapidly and continually increasing. Semantic interoperability is a core problem in such an open and distributed environment [3].

In the description of the OpenGIS service architecture [4] the syntactic and semantic aspects of interoperability are defined as follows: "Syntactical interoperability assures that there is a technical connection, i.e., that the data can be transferred between systems. Semantic interoperability assures that the content is understood in the same way in both systems, including by those humans interacting with the systems in a given context."

In the open and distributed environment of GI web services, the components that are to interoperate are not previously known. The starting point is a requester's spe-

W. Kuhn, M.F. Worboys, and S. Timpf (Eds.): COSIT 2003, LNCS 2825, pp. 186–203, 2003.

cific question rather than a given system. Discovering the services[1] that are appropriate for answering this question from among a large number of available services is a central task within the GI web services domain [5]. Service discovery will therefore be the focus of this paper.

In order to find an appropriate service the requirements resulting from the requester's question have to be matched against descriptions of the the available service implementations. It is in this matchmaking process that semantic interoperability is ensured, making it a crucial part of service discovery (Fig. 1).

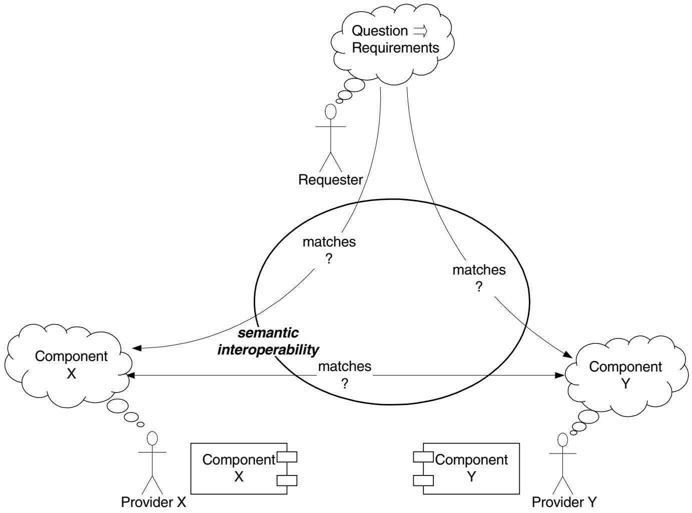

Fig. 1. Semantic interoperability in a GI web services scenario

A large number of languages and technologies have been proposed for service discovery, e.g. for web services in general [6-8], for services and data in the geospatial domain [5, 9, 10], or for software agents and the Semantic Web [11-13][2]. Whether semantic interoperability during service discovery can be achieved in any of these approaches depends on the quality of the information available to the matchmaker on the semantics of requirements and resources. The explicitness, structuring and formality of this information can differ considerably leading to different forms of matchmaking. We are not aware of any framework for classifying the plethora of existing approaches for service discovery with respect to achieving semantic interoperability. Therefore, we propose such a framework in this paper.

The remainder of the paper is structured as follows. In the next section we present several examples of practical problems caused by semantic heterogeneity. The framework for classifying approaches for overcoming semantic heterogeneity is developed in section 0 and applied to the practical problems in section 0. We conclude

---

[1] The notion of *service* in this paper includes both services that can be used to operate on multiple, unspecified datasets (loosely-coupled services) and services that are associated with a specific dataset (tightly-coupled services) [27].

[2] We assume that the reader is familiar with these approaches. Their strengths and weaknesses are outside the scope of this paper and will therefore not be discussed.

the paper by discussing how the results can be applied in the more complex task of (automatic) service composition and by pointing out the next steps for semantic interoperability research along these lines.

2 Examples of Semantic Interoperability Problems

This section presents three examples for problems caused by semantic heterogeneity that we have encountered in our research. They occur in monolithic, at most partially component-based, GIS environments. Nevertheless, they are equally valid for a web service environment.

2.1 Classification of Semantic Heterogeneity

Semantic heterogeneity, the source of semantic interoperability problems, is defined in [14] as the consequence of different conceptualizations and database representations of a real world fact. Two types can be distinguished. *Cognitive heterogeneity* arises when two disciplines have different conceptualizations of real world facts. This becomes a semantic problem when the same names are used for different concepts in both disciplines. Such word pairs are referred to as homonyms. *Naming heterogeneity* refers to different names for identical concepts of real world facts, also called synonyms. The examples subsequently described are classified according to this distinction in order to make sure that both types of heterogeneity are covered.

2.2 Using Topographic Data for Noise Abatement Planning

Situation. To determine which roads could have a considerable noise effect on residential areas those roads touching or crossing residential areas must be identified. German topographic data (Amtliches Topographisch-Kartographisches Informationssystem, ATKIS) contain residential areas and roads as feature classes [15]. Some roads are modeled as lines as shown in Fig. 2 (b).

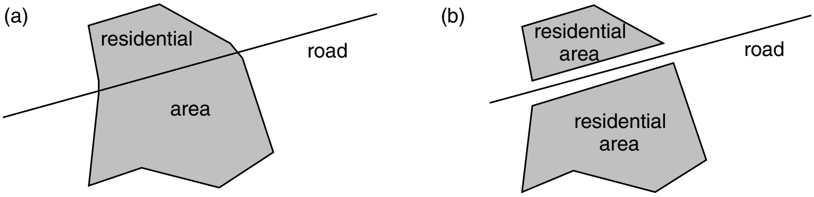

Fig. 2. Different models of roads crossing residential areas

Problem. A user might have the mental concepts of roads and residential areas as depicted in Fig. 2 (a). The system model instead uses representations of roads and residential areas as depicted in Fig. 2 (b). If the user is not aware of the system model (that the terms “residential area” and “road” do not reveal) he might assume that roads overlap residential areas as indicated in Fig. 2 (a) and consequently use the dataset as input for an intersect operation in order to find roads crossing residential areas. However, based on a system model as depicted in Fig. 2 (b) he will not find any roads by doing so, which is correct for the data model of the dataset, but does not meet the user's expectations.

Heterogeneity Type. This example depicts cognitive heterogeneity concerning residential areas; the concepts of user and system regarding the geometric representation are different. The difference is hidden by the homonym "residential area".

2.3 Calculating the Area of Greenland in a Mercator Projection

Situation. In the Mercator map projection features on the reference ellipsoid are projected onto a cylinder touching the equator. This leads to increasing distortion towards the poles and does not preserve areas (see Fig. 3 (b)). For all tasks requiring real world area values, it is not appropriate to calculate the area of features in polar regions, like Greenland, directly from the Mercator projection cylinder (see Fig. 3 (a)).

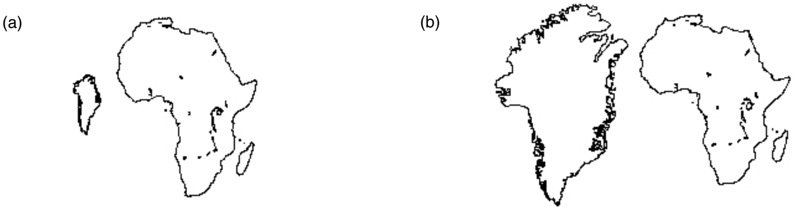

Fig. 3. Greenland and Africa in (a) equal-area Mollweide projection and (b) non-equal-area Mercator projection (images taken from [16])

Problem. Most GIS do not inform users during execution how areas of features are calculated and whether the results reflect the real world area of that feature. The user may expect an area calculation to return the real world area. Such an area calculation would be based on the feature's geometry on the reference ellipsoid. However, if the system's concept of area calculation is based on the feature's geometry on the projection cylinder, the operation will return a completely different result. If the user is not aware of the different concepts of area calculation he will misinterpret the results.

Heterogeneity Type. This example depicts cognitive heterogeneity within the concept of area calculation.

2.4 Topological Operators in GeoMedia and Oracle

This example consists of two parts. First, it describes two operators with the same name and different behavior. Then it describes two operators with different names and equivalent behavior.

Situation. GeoMedia[3] provides a set of topological operators. In addition, it integrates topological operators of the Oracle[4] database system. We look at two GeoMedia operators, called "touch" and "meet", and compare them to an Oracle operator called "touch".

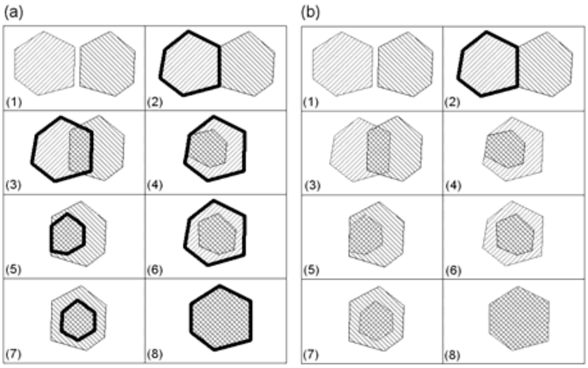

Fig. 4. Regions found (*each marked with a thick line*) by (a) GeoMedia "touch" operator and by (b) Oracle "touch" operator

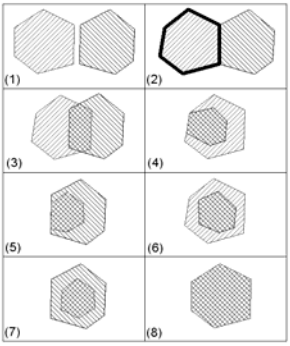

Fig. 5. Region found (*marked with a thick line*) by GeoMedia "meet" operator as well as by Oracle "touch" operator

Problem. We have identified two problems in this example:

1. Although the names are identical, the two "touch" operators of GeoMedia and Oracle return different results (Fig. 4).
2. The GeoMedia "meet" operator and the Oracle "touch" operator, however, find the same regions (Fig. 5 (2)) although they are named differently.

Thus, the names are confusing and misleading, and consequently not useful to the user for deciding if an operation does what he expects.

---

[3] GeoMedia Professional (Intergraph Corp.) V5.0

[4] Oracle 9*i* Release 2 Spatial (Oracle Corp.)

Heterogeneity Type. The first problem is caused by cognitive heterogeneity with the homonym “touch”. The second problem demonstrates naming heterogeneity with the synonyms “meet” and “touch”.

3 Analysis Framework

There are many approaches to ensuring semantic interoperability in the examples presented in the previous section, e.g. [13, 17, 18]. In this section we present a framework for classifying and analyzing such approaches. We define the term *matchmaking* and present different types of matchmaking. We proceed to differentiate several levels of explicitness, structuring and formality for the information required by the matchmaker.

3.1 Matchmaking

In the literature on agent systems matchmaking is defined as mediating among requesters and providers of services for some mutually beneficial cooperation [13]. The process of finding an appropriate service for a certain task can be regarded as matchmaking, too. During matchmaking it is assessed whether (or how well) an available service fits the requirements of the requester.

We distinguish different roles that are played by human actors or system components that exist in the domain of GI web services:

- the *requester* role, which is (ultimately) always played by a human (end user or web service provider),
- the *provider* role, which is also played by a human (web service provider), and
- the *matchmaker* role, which can be played by either a human (one of the above or an independent *broker*) or a matchmaking service.

Note that the same person can take different roles. For example, the person in the requester or provider role can also be responsible for the matchmaking process. As either human or computer can do the matchmaking, two kinds of matchmaking can be distinguished, which represent two endpoints of a continuum:

- *Purely manual matchmaking.* Manual matchmaking is done by a human actor and occurs in the mind of the matchmaker. The matchmaker decides whether or not some service fits the requester’s requirements based on information that is available to him about the service. Manual matchmaking is prone to misunderstandings caused by synonyms and homonyms (section 2.1). In order to mitigate this problem, additional information is collected to reduce ambiguity.
- *Fully automatic matchmaking.* In contrast to manual matchmaking fully automatic matchmaking is always done by a service. This requires *formal* descriptions of requirements and service capabilities. These are matched automatically using an algorithm such as described in [13, 19].

In cases where some of the required formal descriptions are missing, the existing informal descriptions have to be formalized for automatic matchmaking to be applied. Alternatively, the formal descriptions can be made informal and manual matchmaking can be applied. Informalization becomes necessary because formal descriptions are usually difficult to read for non-experts. It should be noted that automatic matchmaking, too, could lead to results unexpected by the requester. This can either be due to explication or formalization errors (i.e. inappropriate capabilities or requirements descriptions) or inappropriate parameterization of the matchmaking algorithms.

3.2 Levels of Explicitness, Structuring and Formality

The quality of the information (metadata) on requirements and service capabilities that is available to the matchmaker is crucial for the matchmaking task. Which information on requirements and on the service has to be made explicit to the matchmaker depends on who does the matchmaking:

- If the *requester does the matchmaking* the requirements are already available in the matchmaker's mind. Therefore, they do not have to be formalized or even made explicit. However, making the requirements explicit and thus reducing ambiguity can help avoiding misinterpretation.
- If the *provider does the matchmaking* the service capabilities are already available in the matchmaker's mind. Therefore, they do not have to be formalized or even made explicit. However, making the capabilities explicit can help to clarify them and discover inconsistencies.
- If an *independent broker does the matchmaking* both requirements and service capabilities have to be made explicit to the matchmaker. A (possibly standardized) structure and formalization might help the broker to do the matchmaking.

The quality of the metadata can vary along three dimensions:

- *Explicitness of information.* The information can be implicit, i.e. only in someone's mind, or explicit, i.e. written down in some language. It is also important to note how complete the available information is, i.e. whether all the information that is required by the matchmaker is available.
- *Structuring of information.* The structure of the information can be implicit and thus unobservable or explicit or even standardized. We refer to the former as unstructured and to the latter as structured information. There are, of course, different levels of structuring [20].
- *Formality of semantics.* The semantics of the concepts used to describe the service can be expressed in ontologies, which are defined as explicit specifications of conceptualizations [21]. A conceptualization is a set of concepts, their definitions and interrelationships [22]. Ontologies can be expressed both informally and formally, i.e using natural or formal languages. There are also intermediate levels of formality [20].

The classification framework could simply consist of these dimensions. However, they are not independent of each other, e.g. the structuring or formality dimensions do not matter if this information is not explicit. Therefore, we propose five levels of explicitness, structuring and formality.

A. *Completely implicit semantics.* The information exists only in the mind of the provider, requester or matchmaker.
B. *Implicit semantics.* Only names (e.g. „forest data“, „web mapping service“) but no metadata are made explicit to refer to services or requirements.
C. *Explicit, unstructured, informal semantics.* Metadata are made explicitly available, but in an unstructured form using natural language text.

D. *Explicit, structured, informal semantics.* Metadata are made explicitly available in a structured form, e.g. referring to metadata standards such as ISO 19115 [23] that – usually informally – specify metadata fields and their semantics. However, with the exception of value lists being specified for some fields, the content of the metadata fields is to be given in the form of free natural language text.

E. *Explicit, formal semantics.* The information is made explicitly available referring to formal ontologies.

These categories are somewhat arbitrary as all three dimensions are continuous. However, we think they represent typical examples for approaches to achieve semantic interoperability. This is illustrated by the scenarios presented in the following section.

3.3 Matchmaking Scenarios

In order to illustrate the levels of explicitness, structuring and formality described in the previous section, three scenarios are depicted. In all of them the requester wants to know the location of forest parcels in the German federal state Northrhine-Westfalia (NRW).

The scenarios represent typical approaches to achieving semantic interoperability at three stages of development. The first scenario shows what is possible and widely practiced by users of the World Wide Web today (levels B and C[5]). The second scenario describes the research and industry attempts made in the GI community (level D), most notably in the OpenGIS Consortium (*http://www.opengis.org*) and the ISO Technical Committee 211 (*http://www.isotc211.org*). The last scenario presents ideas that are currently discussed in the Semantic Web and agent systems communities (level E).

Note that the role labeled *requester* in the following figures could either represent an end user who wants an appropriate service to answer his question, or a service provider who wants to find appropriate services to build a complex service that performs a specific task. The actor or component responsible for the matchmaking is highlighted in gray.

Scenario 1 – Manual Matchmaking Based on Names or Unstructured and Informal Information. In the first scenario (Fig. 6) the capabilities of the services are not made explicit by their providers. The only clues for the requester to what the services are doing or which data they provide are their names. One means for finding appropriate services by their name is through a keyword search in an Internet search engine like Google. In such a search the requester can encounter the following problems:

- *No match.* Services that fit the requester's requirements are not found at all, because their names do not match the keywords included in the requester's query. The simplest reason for this are spelling differences or mistakes. Leaving these aside, the problem can be classified as a case of naming heterogeneity (section 2.1): The conceptualizations of requester and provider are sufficiently similar for the task at hand but concepts are given different names (synonyms). This can have

---

[5] Level A is not considered because service discovery becomes extremely difficult or even impossible when the semantics of requirements or capabilities are completely implicit.

several reasons. Either the name of the service or the keywords used in the query are not appropriate, i.e. they do not well reflect the service capabilities or the requester requirements, respectively. Or both keywords and names are appropriate (within their respective domain), but requester and providers belong to different information communities.

- *Unsuitable match.* Services that are found because their name matches the keywords included in the requester's query do not fit the requester's requirements. The conceptualizations of requester and provider are different but are given the same names (homonyms). This case can be classified as cognitive heterogeneity leading to naming conflicts (section 2.1). The possible reasons for this can again be inappropriate names or differing information communities.

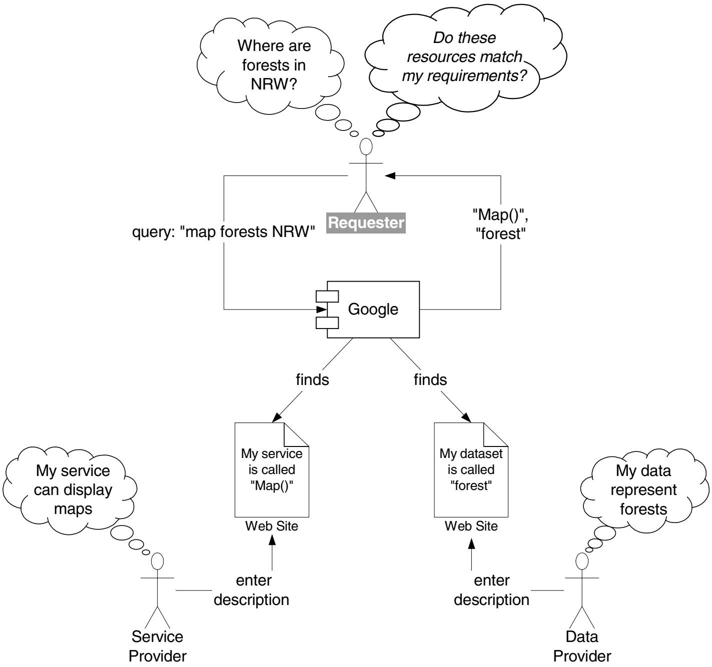

Fig. 6. Manual matchmaking based on names (level B) or unstructured and informal information (level C)

An explicit requirements specification can help the requester to do the matchmaking, because the process of explication often helps to clarify and disambiguate ideas on requirements. Also explicitly describing the service's capabilities rather than only giving a name can improve the matchmaking by reducing guesswork. This is the case of explicit, but unstructured and informal description of semantics. However, misinterpretation is still possible if the descriptions are ambiguous or incomplete. These two cases are currently the most common ones as service descriptions are either informal or missing completely.

Scenario 2 – Manual Matchmaking with Standardized Metadata. In the second scenario (Fig. 7) the providers' conceptualizations are made explicit and are recorded in metadata documents whose structure is well known and which are made available through one (or several) registries. The requester can search a registry using keywords for all of the fields provided by its query interface. He can then use the returned metadata documents to assess whether or not the services fit his requirements. He might need to access other documents that the metadata documents refer to, e.g. a feature type catalog providing definitions for feature classes or ISO standards defining units of measurement.

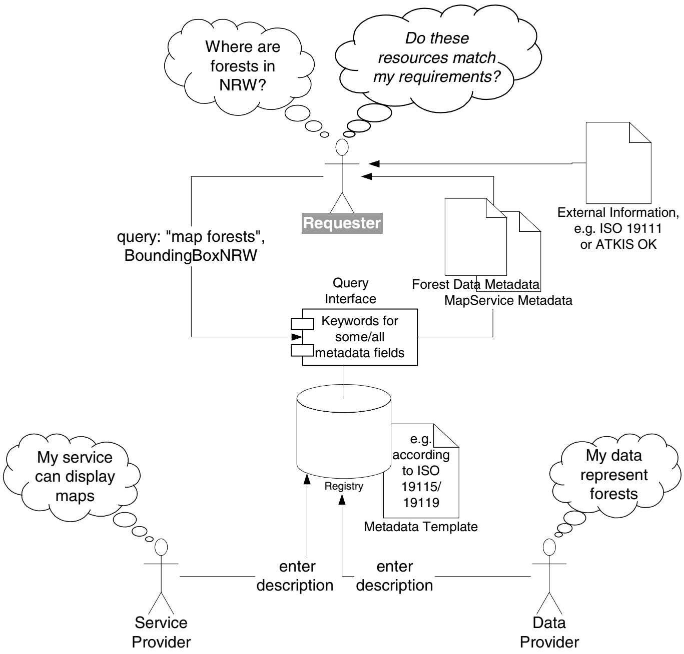

Fig. 7. Manual matchmaking with standardized metadata, e.g. ISO 19115/19119 (level D)

As the matchmaking in this scenario is still based on keywords, the problems described for the first scenario can still occur. There can be ambiguity in either the metadata entries themselves or in the referenced documents (e.g. the feature catalogue). However, this can be considerably reduced by using standardized documents, by providing a controlled vocabulary (e.g. lists of keywords), and by referring to other standardized or at least widely known and agreed-upon documents.

Scenario 3 – Automatic Matchmaking with Formal Metadata. In the last scenario (Fig. 8) the conceptualizations of requester and providers are not only explicit but also formalized. They use concepts from existing *domain ontologies* [24] to formulate their requirements or advertisements, respectively. A service automatically matches the requester's requirements against advertisements stored in its registry using a matchmaking algorithm such as described in [13].

It is assumed that by using formal descriptions of semantics and automatic matchmaking algorithms problems such as those described in the previous scenarios can be avoided [12, 25]. However, in this scenario, too, problems similar as those identified in the previous scenarios, albeit for different reasons, can occur.

- *No match.* Services that fit the requester's requirements are not found at all because the matchmaking algorithm is too rigorous. In [13] a threshold value has to be specified by the requester indicating which degree of similarity between advertisements and requirements is still acceptable.
- *Unsuitable match.* Services that are found do not fit the requester's requirements. This, too, can be caused by the calibration of the matchmaking algorithm. Here, the matchmaking algorithm is too tolerant because the threshold value is too low. Another possible reason is that the requirements document does not correctly reflect the requester's requirements or the capabilities documents do not correctly reflect the providers' conceptualization of the service. We refer to these kinds of errors as *explication* or *formalization errors*, respectively.

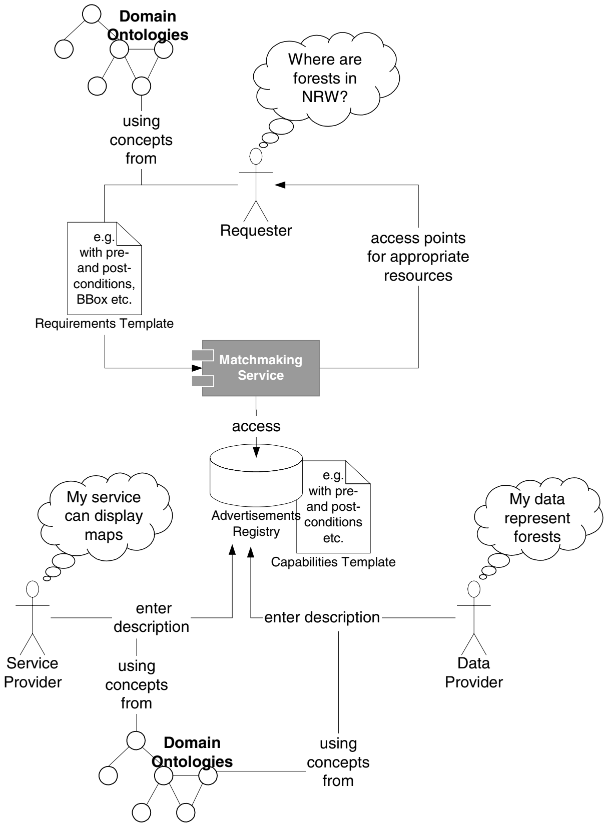

Fig. 8. Automatic matchmaking with formal metadata (level E)

3.4 Likelihood of Misunderstanding

Misunderstandings can occur in all matchmaking scenarios described in the previous sections. Summarizing the arguments from the previous section, Table 1 gives an estimate of the likelihood of misunderstandings for all possible combinations of explicitness, structuring and formality levels described above. It is assumed that the requester does the matchmaking.

Table 1. Likelihood of misunderstandings for different levels of requirements and capabilities descriptions if the requester does the matchmaking. Shading: white – manual matchmaking possible, light gray – manual matchmaking possible but difficult for non-experts, dark gray – automatic matchmaking possible. The scenarios described in the previous section are framed

requirements service capabilities	(completely) implicit	explicit, unstructured, informal semantics	explicit, structured, informal semantics	explicit, formal semantics
completely implicit	matchmaking impossible			
implicit	**I** highly likely	highly likely	highly likely	highly likely
explicit, unstructured, informal semantics	highly likely	likely	likely	likely
explicit, structured, informal semantics	likely	possible	**II** possible	possible
			possible (automatic matchmaking limited)	
explicit, formal semantics	likely	possible	possible	unlikely
			possible (autom. matchm. limited)	**III** unlikely

4 Analysis of Examples

After having presented the framework for classifying matchmaking approaches we show in this section how it can be applied to the examples presented in section 0. In the following tables the first row lists the information required by the matchmaker in order to find resources appropriate for answering the requester's question. The names of the concepts appear in italics. The remaining rows contain an analysis of the availability, quality and source of the information in each of the three scenarios presented in section 0.

4.1 Using Topographic Data for Noise Abatement Planning

This example depicts the requester's attempt to intersect residential areas with roads. This involves a matchmaking process for which information about the requester's conceptualization of *road*, *residential area* and the operators *touch* and *cross* as well as information about the ATKIS geometry model are needed (Table 2).

Table 2. Application of the classification framework to example 1 – Using topographic data for noise abatement planning. (The table is split into two for enhanced readability)

		information required by the matchmaker				
		requester conceptualization of *road* and *residential area*	requester conceptualization of *touch* and *cross*	ATKIS geometry model for *road* and *residential area*		process model for geoprocessing operations, e.g. *intersect* or *buffer*
scenario 1	available	✓	✓	–	✓	✓
	level	implicit	implicit	n.a.	implicit	implicit
	source	requester's mind	requester's mind	n.a.	requester's mind (if he is an ATKIS expert) or dataset (accessible via visualization of data, requires GIS expertise)	requester's mind (if he is an expert of the specific GIS) or trial and error (requires GIS expertise)
scenario 2	available	✓	✓	✓		✓
	level	implicit	implicit	explicit, structured, informal		explicit, structured, informal
	source	requester's mind	requester's mind	The ISO metadata standard supports references to external feature type catalogs like that of ATKIS as well as graphic overviews [23].		The ISO services standard provides a template for describing services [4]. Alternatives are UDDI [7], WSDL [8], Capabilities XML [26]. They focus on operation signatures; descriptions are available only on service level and appear as free text. ISO in addition provides free text descriptions on the operation level. The ISO spatial schema standard provides information for filling such a template [27]. They consist of free text descriptions and formalized operation signatures.
scenario 3	available	✓	✓	✓		✓
	level	explicit, formal	explicit, formal	explicit, formal		explicit, formal
	source	domain ontology chosen by the requester to describe his task	domain ontology chosen by the requester to describe his task	domain ontology based on ATKIS feature type catalog [15]		(geo)processing domain ontology, e.g. based on ISO spatial schema standard [27]

In scenario 1 the intersection attempt will only be successful if the requester is an expert who is aware of how the ATKIS geometry model will fit his requirements. In scenario 2 the intersection attempt will be successful if the requester is willing to spend the time to access and understand the available metadata and perform the matchmaking manually. In scenario 3 the intersection attempt will be valid even if the requester is no ATKIS expert, because the information needed for the matchmaking is available in formal and explicit form, making automatic matchmaking possible. The result of the matchmaking process may be that the intersection is not possible because the mapping from system to requester concepts would require additional services that are not available. Nevertheless, even in this case the requester is saved from misinterpreting the results of the intersection.

4.2 Calculating the Area of Greenland in a Mercator Projection

This example depicts the requester's attempt to calculate the real world area of Greenland displayed with a GIS using the Mercator projection. This involves a matchmaking process for which information about the requester's conceptualization of area calculation, the system model of area calculation and indirectly information about the attributes of Mercator projections is needed (Table 3).

In scenario 1 the area calculation attempt is likely to lead to misinterpretation as long as the requester is no GI expert. In scenario 2 the area calculation attempt is likely to be canceled. If the requester is willing to spend the time to access and understand the available metadata, he becomes aware of that the calculated area will not meet his requirements of representing the real world area. However, in scenario 2 no further solution is offered. In contrast, in scenario 3 the area calculation attempt may be successful, because all information needed for the matchmaking process is available in formal and explicit form. The requester is made aware of that his requirements differ from the system's abilities. It might be possible to search for a service that is able to calculate the area according to the requester's requirements. In this scenario the requester does not need any knowledge about projections and area calculation operations.

4.3 Topological Operators in GeoMedia and Oracle

This example depicts the requester's attempt to find operations that return geometry features whose boundaries intersect but whose interiors do not. To find the appropriate operations the requester's requirements have to be matched with the systems' capabilities. For the matchmaking process information about the requester's conceptualization of *touch* is needed as well as the process models of the available operations of the systems, in this case GeoMedia and Oracle (Table 4).

In scenario1 the attempt to find the appropriate operation is likely to lead to misinterpretations if the requester is no system expert. In scenario 2 the attempt may be successful if the requester is willing to spend the time to access and understand the available metadata. He then will learn about the meaning of the different operations and will be able to perform a manual matchmaking. In scenario 3 the matchmaking will be successful. Using terms from a domain ontology the requester can specify his requirements formally and explicitly. Based on this formal and explicit specification the appropriate operations can be chosen automatically from among the available operations.

Table 3. Application of the classification framework to example 2 – Calculating the area of Greenland in a Mercator projection

		information required by the matchmaker				
		requester conceptualization of *area calculation*	system model of *area calculation* (possibly including attributes of the projection, see next row)		attributes of *Mercator projection*	
scenario 1	available	✓	–	✓	–	✓
	level	implicit	n.a.	implicit	n.a.	implicit
	source	requester's mind	n.a.	requester's mind (if he is a GI expert)	n.a.	requester's mind (if he is a GI expert)
scenario 2	available	✓	✓		✓	
	level	implicit	explicit, structured, informal		explicit, structured, informal	
	source	requester's mind	The operation signatures can be described in the same way as for the intersect and buffer operations in Table 2. The ISO metadata standard [23] provides attributes for operations which can be applied to the dataset. However, the requester has to judge whether the results (e.g. area calculation) fit his expectations. (see same column next row).		The ISO standard for spatial referencing by coordinates provides a free text description indicating for which application a coordinate reference system is valid [28].	
scenario 3	available	✓	✓		✓	
	level	explicit, formal	explicit, formal		explicit, formal	
	source	domain ontology chosen by the requester to describe his task	(geo)processing domain ontology, e.g. based on ISO spatial schema standard [27]		domain ontology for projections, e.g. based on ISO standard for spatial referencing by coordinates [28]	

5 Conclusions and Future Work

We have presented a framework for classifying approaches to achieving semantic interoperability in the domain of GI web services. The framework focuses on the process of matchmaking as this is where semantic interoperability is ensured. Therefore approaches to achieving semantic interoperability are classified according to the quality of the information that is available to the matchmaker.

The application of the framework has been illustrated by analyzing existing approaches to solving examples of semantic interoperability problems. In scenario 1 misinterpretations are likely to occur unless the requester is an expert for the components employed. In scenario 2 misinterpretations are less likely if the requester is willing to spend the time to access and understand the available metadata. In scenario 3 misinterpretations are unlikely, even for non-experts, as automatic matchmaking is applied. However, there is still the possibility that the services required for the requester's query are not available.

Table 4. Application of the classification framework to example 3 – Topological operators in GeoMedia and Oracle

<table>
<tr><th colspan="2" rowspan="2"></th><th colspan="3">information required by the matchmaker</th></tr>
<tr><th>requester conceptualization of *touch*</th><th colspan="2">process models for *touch* operations (GeoMedia and Oracle) and *meet* operation *(GeoMedia)*</th></tr>
<tr><td rowspan="3">**scenario 1**</td><td>available</td><td>✓</td><td>–</td><td>✓</td></tr>
<tr><td>level</td><td>implicit</td><td>n.a.</td><td>implicit</td></tr>
<tr><td>source</td><td>requester's mind</td><td>n.a.</td><td>requester's mind (if he is an expert of the specific GIS) or trial and error (requires GIS expertise)</td></tr>
<tr><td rowspan="3">**scenario 2**</td><td>available</td><td>✓</td><td colspan="2">✓</td></tr>
<tr><td>level</td><td>implicit</td><td colspan="2">explicit, structured, informal</td></tr>
<tr><td>source</td><td>requester's mind</td><td colspan="2">The ISO services standard provides a template for describing services [4]. The ISO spatial schema standard provides information for filling such a template [27]. See also intersect and buffer operations in Table 2.</td></tr>
<tr><td rowspan="3">**scenario 3**</td><td>available</td><td>✓</td><td colspan="2">✓</td></tr>
<tr><td>level</td><td>explicit, formal</td><td colspan="2">explicit, formal</td></tr>
<tr><td>source</td><td>domain ontology chosen by the requester to describe his task</td><td colspan="2">(geo)processing domain ontology, e.g. based on ISO spatial schema standard [27]</td></tr>
</table>

The analysis of practical problems only presents a first application of the framework. We believe the framework to be valuable to the GI research community for structuring the domain of semantic interoperability research, because it supports the following tasks:

- The information required for the matchmaking process can be identified.
- The required information can be classified according to the qualities explicitness, structuring and formality.
- It can be assessed which quality level of the required information is appropriate for the task at hand.
- The different levels of explicitness, structuring and formality can easily be associated to predefined scenarios that indicate possible implementation methods .
- In the combination of the above reasons, researchera can classify their approach and judge whether the applied methods are appropriate for the task at hand.

Future work must look at the role that service discovery plays within the larger task of service composition. It will also be examined whether other sub-tasks play a role in ensuring semantic interoperability in (especially ad-hoc) service composition. For this an abstract model of service composition should be developed. Such a model could be valuable for the standardization efforts in OGC and ISO TC 211, where the task of service composition has not yet been thoroughly explored.

It also remains an open question whether examples like those presented in this paper represent a specific (i.e. spatial) kind of semantic heterogeneity or whether they can be treated in the same way as other (non-spatial) semantic problems. If the latter turns out to be possible the framework should be adjusted accordingly.

Acknowledgements

Comments from Werner Kuhn to earlier drafts of this paper helped clarify the ideas. The work presented in this paper has been partially supported by the German Ministry for Education and Science as part of the GEOTECHNOLOGIEN program (grant number 03F0369A) and can be referenced as publication no. GEOTECH-23. Furthermore, support from the European Commission through the ACE-GIS (grant number IST-2002-37724) and BRIDGE-IT (grant number IST-2001-34386) projects are gratefully acknowledged.

References

1. Abel, D. J., Gaede, V. J., Taylor, K. L., Zhou, X.: SMART: Towards Spatial Internet Marketplaces. Geoinformatica 3 (1999) 141-164
2. Groot, R., McLaughin, J.: Geospatial data infrastructure – Concepts, cases, and good practice. Oxford University Press (2000)
3. OGC: OpenGIS Web Services Architecture. OpenGIS Consortium, OpenGIS Discussion Paper OGC 03-025 (2003)
4. ISO/TC-211, OGC: Geogaphic information – Services (ISO/DIS 19119) v4.3. International Organization for Standardization & OpenGIS Consortium (2002)
5. Egenhofer, M.: Toward the Semantic Geospatial Web. In: Proc. The 10th ACM International Symposium on Advances in Geographic Information Systems (ACM-GIS) (2002)
6. OASIS: OASIS/ebXML Registry Services Specification v2.5. OASIS/ebXML Registry Technical Committee (2003)
7. Bellwood, T., Clément, L., Ehnebuske, D., Hately, A., Hondo, M., Husband, Y. L., Januszewski, K., Lee, S., McKee, B., Munter, J., von Riegen, C.: UDDI v 3.0. (2002)
8. Chinnici, R., Gudgin, M., Moreau, J.-J., Weerawarana, S.: Web Services Description Language (WSDL) v1.2. (2002)
9. Reed, C., Nebert, D.: The Importance of Catalogs to the Spatial Web. An OGC White Paper. (2002)
10. OGC: OWS1.2 UDDI Experiment. OpenGIS Consortium, OGC 03-028 (2003)
11. Constantinescu, I., Faltings, B.: Efficient Matchmaking and Directory Services. Swiss Federal Institute of Technology, Techn. Report IC/2002/77 Lausanne, Switzerland (2002)
12. Paolucci, M., Kawamura, T., Payne, T. R., Sycara, K.: Semantic Matching of Web Service Capabilities. In: Proc. 1st International Semantic Web Conference (ISWC2002) (2002) 333-347
13. Sycara, K., Widoff, S., Klusch, M., Lu, J.: Larks: Dynamic Matchmaking Among Heterogeneous Software Agents in Cyberspace. In: Proc. First International Joint Conference on Autonomous Agents and Multi-Agent Systems (2002) 173-203
14. Bishr, Y.: Overcoming the semantic and other barriers to GIS interoperability. International Journal of Geographical Information Science 12 (1998) 299-314

15. AdV-Arbeitsgruppe ATKIS: ATKIS-Objektartenkatalog Basis-DLM. (2002)
16. Furuti, C. A. Useful Map Properties.[Online]. Available: http://www.progonos.com/furuti/MapProj/Normal/CartProp/
17. Visser, U., Stuckenschmidt, H.: Interoperability in GIS - Enabling Technologies. In: Proc. 5th AGILE Conference on Geographic Information Science (2002) 291-297
18. Kuhn, W., Raubal, M.: Implementing Semantic Reference Systems. In: Proc. 6th AGILE Conference on Geographic Information Science (2003)
19. Sycara, K., Klusch, M., Widoff, S., Lu, J.: Dynamic Service Matchmaking Among Agents in Open Information Environments. ACM SIGMOD Record 28 (1999) 47-53
20. Uschold, M.: Knowledge level modelling: concepts and terminology. The Knowledge Engineering Review 13 (1998) 5-29
21. Gruber, T. R.: Toward Principles for the Design of Ontologies Used for Knowledge Sharing. International Journal of Human-Computer Studies 43 (1995) 907-928
22. Uschold, M., Gruninger, M.: Ontologies: Principles, Methods and Applications. The Knowledge Engineering Review 11 (1996) 93-136
23. ISO/TC-211: Geogaphic information – Metadata (ISO/FDIS 19115). International Organization for Standardization (2003)
24. Guarino, N.: Formal Ontology and Information Systems. In: Proc. Formal Ontology in Information Systems (FOIS'98) (1998) 3-15
25. Guarino, N.: Semantic Matching: Formal Ontological Distinctions for Information Organization, Extraction, and Integration. In: Proc. Information Extraction: A Multidisciplinary Approach to an Emerging Information Technology, International Summer School (SCIE-97) (1997) 139-170
26. de La Beaujardière, J.: Web Map Service Implementation Specification. Open GIS Consortium (2002) 82
27. ISO/TC-211: Geogaphic information – Spatial Schema (ISO/DIS 19107). International Organization for Standardization (2002)
28. ISO/TC-211: Text for FDIS 19111 Geogaphic information - Spatial Referencing by Coordinates. Final Draft Version. International Organization for Standardization, (2002)