DGK Deutsche Geodätische Kommission
der Bayerischen Akademie der Wissenschaften

Reihe C | Dissertationen | Heft Nr. 732

Christian Rüh

Marine spatial data infrastructures
– Approaches on evaluation, design and implementation –

München 2014

Verlag der Bayerischen Akademie der Wissenschaften
in Kommission beim Verlag C. H. Beck

ISSN 0065-5325 ISBN 978-3-7696-5144-7

Deutsche Geodätische Kommission
der Bayerischen Akademie der Wissenschaften

Reihe C | Dissertationen | Heft Nr. 732

Marine spatial data infrastructures
– Approaches on evaluation, design and implementation –

Dissertation
to obtain the degree of
doctorate in engineering (Dr.-Ing.)
at the Faculty of Agricultural and Environmental Sciences
of the University of Rostock

submitted by

MSc Christian Rüh
born on 17.05.1985 in Pasewalk

München 2014

Verlag der Bayerischen Akademie der Wissenschaften
in Kommission beim Verlag C. H. Beck

ISSN 0065-5325 ISBN 978-3-7696-5144-7

Adresse der Deutschen Geodätischen Kommission:

Deutsche Geodätische Kommission

Alfons-Goppel-Straße 11 ● D – 80 539 München

Telefon +49 – 89 – 23 031 1113 ● Telefax +49 – 89 – 23 031 - 1283 / - 1100

e-mail hornik@dgfi.badw.de ● http://www.dgk.badw.de

Advisors: Prof. Dr.-Ing. Ralf Bill, University of Rostock
Dr.-Ing. Rainer Lehfeldt, Federal Waterways Enineering and Research Institute (BAW), Hamburg
Prof. Dr. Christophe Claramunt, Naval Academy Research Institute, France

Date of submission: 21.02.2014
Date of viva: 02.06.2014

Diese Dissertation ist auf dem Server der Deutschen Geodätischen Kommission unter <http://dgk.badw.de/>
sowie auf dem Server der Universität Rostock unter <urn:nbn:de:gbv:28-diss2014-0094-3> elektronisch publiziert

---

ISSN 0065-5325 ISBN 978-3-7696-5144-7

Abstract

Humanity is reliant on and attracted to the marine environment. In order to make use of its resources and to protect it guidelines and directives were and are being developed. To aid the implementation of such efforts spatial data infrastructures (SDIs) can be used. These allow administrative officers and scientists inter alia to publish data and prepare reports. They can also be used by the public or contribute to political decision-making processes. In the marine domain such SDIs are called marine spatial data infrastructures (MSDIs) and Germany began developing one – called MDI-DE – in 2010.

Other countries developed MSDIs well before the year 2010 which opens up the opportunity to learn from these approaches. In order to have a rather objective and comparable base an evaluation framework is needed. This implies equal procedures for each MSDI which means that one cannot lose track of things. Furthermore this indicates that the results of the evaluations elaborate the pros and cons (potential pitfalls and things done well). This thesis develops such an evaluation framework to assess MSDIs and applies it to the MSDIs of Ireland, the UK, the USA, Canada and Australia.

Another opportunity that opened up because Germany is building a MSDI for the first time is that its development can be based on and guided by a reference model. A reference model structures large and complex distributed systems such as MSDIs with the help of several viewpoints respectively submodels. These allow focusing on specific parts of an architecture and are necessary because different stakeholders have different interests in such a system. The reference model this thesis proposes consists of five such submodels: business, role, process, architecture and implementation.

The reference model inter alia envisaged setting up infrastructure nodes with distributed services. Services are a base of a SDI to work. They are also required by the INSPIRE directive. INSPIRE proposes requirements regarding performance so that services are conveniently accessible. Furthermore INSPIRE requires data and metadata to follow a specific structure. The same is true for services themselves because they have to follow given standards and specifications e.g. by the International Organization for Standardization (ISO) and the Open Geospatial Consortium (OGC). This thesis uses existing tools to monitor and evaluate services and attempts to clarify whether the results of the tools are comparable and if the INSPIRE requirements can be evaluated with such tools.

Lastly, an important aspect of MSDIs, in particular, are terms which are combined in so-called controlled vocabularies respectively thesauri because MSDIs are more scientifically oriented and interdisciplinary than terrestrial SDIs. The existing vocabularies did not allow to be used by systems (e.g. for metadata annotation) or be maintained by marine experts (e.g. by using a web authoring tool). To allow such usages this thesis implements a tool to convert the vocabularies into the Simple Knowledge Organisation System (SKOS) format. The conversion into SKOS allows importing the vocabularies into an online thesaurus management tool.

Altogether this thesis with focus on specific aspects of evaluation, design and implementation of marine spatial data infrastructures should scientifically support the development of the German approach for the MDI-DE.

Keywords:
Spatial data infrastructure, marine, INSPIRE, reference model, services, modelling

Zusammenfassung

Die Menschheit ist sowohl von der Meeresumwelt angezogen, als auch auf diese angewiesen. Um sie zu schützen und ihre Ressourcen zu nutzen, wurden und werden Richtlinien entwickelt, deren Anforderungen u. a. mit Geodateninfrastrukturen (GDIen) erfüllt werden können. Diese ermöglichen u. a. verantwortlichen Sachbearbeitern oder Wissenschaftlern Daten zu veröffentlichen und Berichte zu erstellen; können aber auch von der Öffentlichkeit verwandt werden oder zu politischen Entscheidungsprozessen beitragen. Im marinen Umfeld heißen solche GDIen marine Geodateninfrastrukturen (MGDIen) und Deutschland entwickelt eine solche – genannt MDI-DE – seit 2010.

Andere Länder entwickelten MGDIen bereits weit vor dem Jahr 2010, was die Möglichkeit eröffnete, von diesen Ansätzen zu lernen. Um eine relativ objektive Basis zu haben, benötigt man einen Bewertungsrahmen. Dieser ermöglicht es, bei der Analyse der MGDIen stets gleich vorzugehen und somit im Ergebnis der Bewertungen die Vor- und Nachteile der existierenden Ansätze herausarbeiten. Diese Arbeit konzipiert einen solchen Bewertungsrahmen für MGDIen und wendet diesen auf die MGDIen von Irland, Großbritannien, USA, Kanada und Australien an.

Eine weitere Chance, die sich dadurch ergibt, dass Deutschland zum ersten Mal eine MGDI aufbaut, ist die Möglichkeit, diese auf Grundlage eines Referenzmodells zu entwickeln. Ein Referenzmodell erlaubt die Strukturierung großer und komplexer verteilter Systeme, wie z. B. MGDIen, mithilfe mehrerer Teilmodelle. Diese ermöglichen u. a. die Konzentration auf bestimmte Teile einer Architektur. Das Referenzmodell, das in dieser Arbeit aufgebaut wird, gliedert sich in die Teilmodelle Geschäfts-, Rollen-, Prozess-, Architektur- und Implementierungsmodell.

Das Referenzmodell sieht u. a. die Einrichtung von Infrastrukturknoten mit Diensten vor. Dienste werden auch von der INSPIRE-Richtlinie gefordert, die überdies Anforderungen in Bezug auf die Leistungsfähigkeit von Diensten definiert. Darüber hinaus müssen Dienste vorgegebenen Standards und Spezifikationen der International Organization for Standardization und des Open Geospatial Consortiums entsprechen. Diese Arbeit stützt sich auf bestehende Werkzeuge zur Überwachung und Bewertung von Diensten und untersucht die Vergleichbarkeit der Ergebnisse der Werkzeuge und ob die Anforderungen von INSPIRE mit solchen Werkzeugen bewertet werden können.

Abschließend sind Begriffe, die in sogenannten kontrollierten Vokabularen beziehungsweise Thesauri zusammengefasst werden, ein wichtiger Aspekt insbesondere von MGDIen, da MGDIen in höherem Maße wissenschaftlich orientiert und fachübergreifender sind als terrestrische GDIen. Mit den vorhandenen Vokabularen war es nicht möglich, sie von Systemen (z. B. für die Beschreibung von Metadaten) verwenden oder sie von Wissenschaftlern gemeinschaftlich pflegen zu lassen. Um solche Nutzungen zu ermöglichen, wird in dieser Arbeit ein Werkzeug entwickelt, das Vokabulare in das Simple Knowledge Organisation System (SKOS) Format konvertiert, was den Import der Thesauri in ein Web-Thesaurus-Management-Tool erlaubt.

Insgesamt soll diese Arbeit über bestimmte Aspekte der Bewertung, des Entwurfs und der Umsetzung von Marinen Dateninfrastrukturen die Entwicklung des deutschen Ansatzes für die MDI-DE wissenschaftlich unterstützen.

Schlagwörter:
Geodateninfrastruktur, marin, INSPIRE, Referenzmodell, Dienste, Modellierung

Contents

1 Introduction

> *"Humankind is extremely reliant on the oceans, as a source of food and raw materials, as a climate regulator, for transportation, for disposal of waste products, and for recreation."*
>
> (Strain et al., 2006, p. 431)

Marine environments are very important to mankind because resources can be exploited, habitats can be found and industries can produce goods that can then be shipped over the seas. Furthermore especially coastal environments are valued living spaces and recreational areas. These are reasons why *"[. . .] half the world's population lives within 60 km of the shoreline [. . .] "*[1].

However, the negative side to this are *"[. . .] environmental modification and deterioration through landfill, dredging, and pollution caused by urban, industrial, aquaculture and agricultural activities."*[2] This means that the marine environment is in danger and foresighted management and actions are needed. Such actions are derived inter alia from directives that demand continuous monitoring efforts and periodic reports.

A spatial data infrastructure (SDI) can support the fulfillment of directives' requirements. It can help administrative officers and scientists to find data they need, publish data so that other users are able to use it and prepare reports that reflect the state of marine environments. Decisions can be made by politicians, environmental agencies and so on based on these reports and the data a SDI makes available. Furthermore a SDI is an instrument to inform the public.

1.1 Motivation

Germany, in contrast to other countries, did not have a marine spatial data infrastructure (MSDI) until 2014. To make data access easier and merge information concerning different topics – such as coastal engineering, hydrography and surveying, protection of the marine environment, maritime conservation, regional planning and coastal research – the Federal Ministry of Education and Research (BMBF) funded the project MDI-DE[3] in order to develop a MSDI for Germany. Easier data access should support institutions and authorities in their daily work because it is easier for employees and/or scientists to find the data they need.

Apart from easier data access and merging information through a central geoportal *reporting* to specific marine directives is an relevant aspect and will be even more important in the future when the directives are implemented and require data and reports on a specific time cycle. On the European level Germany has to report to the INSPIRE[4] (Infrastructure for Spatial Information in the European Community) directive as well as the Marine Strategy Framework Directive[5] (MSFD), the Water Framework Directive[6] (WFD) and Natura2000 with their regulation counterparts for Germany and its federal states (Meeresstrategie-Rahmenrichtlinie [MSRL], Wasserrahmenrichtlinie [WRRL], Fauna-Flora-Habitat-Richtlinie [FFH-RL], Vogelschutzrichtlinie [VS-RL]). A central geoportal will help to comply with the reporting requirements because of its standardization respectively harmonization and the easy and centralized data access. Furthermore as the development of a MSDI

[1] (Bartlett et al., 2004, p. 2)
[2] (Bartlett et al., 2004, pp. 2)
[3] www.mdi-de.org
[4] http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32007L0002:EN:NOT
[5] http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32008L0056:EN:NOT
[6] http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32000L0060:EN:NOT

brings together marine experts it can be defined and implemented how to report to the directives (e.g. data harmonization is needed so that biological and chemical parameters are represented in a comparable way).

Data harmonization does also require metadata harmonization. Not only for data and metadata but for many aspects terms are of great importance. Terms are created and merged into *thesauri* by marine experts so that terms are clear and non-ambiguous. Since Germany has not had a MSDI so far there are several marine thesauri – developed by several institutions – in existence which means that terms may be included several times and that definitions may vary and so on.

1.2 Objectives

The previous section stated aspects of what is missing respectively aspects that are worth working on. The objectives presented in this section derive from these aspects and implement the motivated aspects. It has to be noted that this thesis can only implement respectively design selected aspects of an (M)SDI.

MSDI from scratch
Building an initial MSDI in Germany is challenging as well as promising. On the one hand it is challenging because of all the coordination and effort that have to be put into such a development. On the other hand it is promising because standards have matured, state-of-the-art technologies can be used and because other countries already have been working on MSDIs as well. Therefore the first step is to look for pre-existing MSDIs enabling learning from them. The second step is to analyze them. However, there is an intermediate step between these two because firstly it has to be defined how to analyze them to be able to compare them which might make learning from them easier. The final step is to extract potential pitfalls on the one hand and advantages respectively good ideas otherwise.

Easier and central data access
Apart from learning from other infrastructure initiatives an early step in (M)SDI development is to become clear of what is already there. A SDI brings together many actors (authorities, institutions etc.) and therefore much data and metadata (stored in so called *infrastructure nodes*). Before the development begins all the data and metadata sets have to be known to see how and if they fit into the infrastructure. If the data and metadata is not already available through services an early step in SDI development is setting up services so that files do not have to be transferred and that the most up-to-date version of data and metadata is available from *one* source every stakeholder has access to.

Reporting
After the data is made available easily to every stakeholder through services the availability of services enables actors to use the data, for instance, to comply with reporting requirements. For this usage on the one hand aspects like data modelling (i.e. what the data have to look like) and other formal requirements concerning data and metadata are important. On the other hand – if data have to be published respectively transmitted via services – there also are requirements concerning service quality because not only there are user expectations regarding availability and performance but also reports that have to be prepared based on the services which means that the services respond in an acceptable time frame and are permanently available.

Thesauri harmonization
For reporting – but also for other tasks such as metadata annotation and search – terms are important because they have to be clearly defined so that everybody can be certain what is meant with a specific term. These terms come from thesauri. The first task is to identify the thesauri handling marine terms in Germany. If there is more than one thesaurus in existence the thesauri have to be harmonized so that terms are not defined multiple times, especially that terms are not defined differently. To be able to use thesauri – or *the* harmonized thesaurus – for other tasks and in particular to ease the harmonization process (through the use of an editorial system) a web-based thesauri management system is used into which the thesauri are imported first.

1.3 Outlook on the thesis

Figure 1.1 gives an overview of this thesis and shows how the chapters and sections are related to one another.

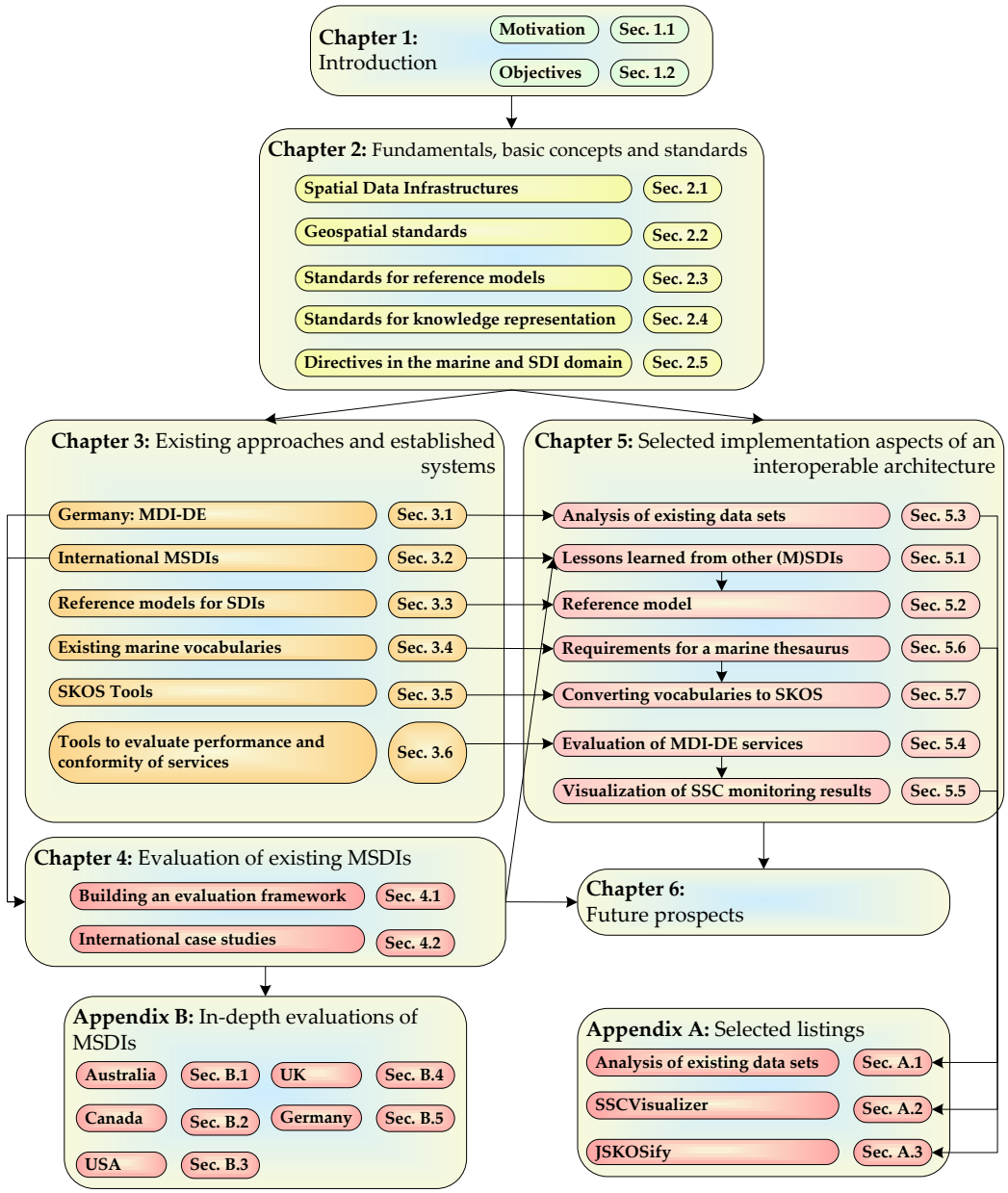


Figure 1.1: *Chapters of this thesis and their relationships*

Chapter 2 – Fundamentals, basic concepts and standards Based on the motivation of this thesis and its objectives there is a range of technologies, standards and concepts that are needed to comply with the objectives. Because MDI-DE is another spatial data infrastructure (SDI) it is important to know its components, what interoperability means and what the unique features of marine SDIs (in comparison to terrestrial SDIs) are. Since SDIs are based on services and would not be possible without these standards for services are stated in conjunction with standards for the underlying data and metadata. MDI-DE is a development from scratch based on these standards. That means that a reference model could be used to aid and support the development of MDI-DE which is why standards for reference models are also presented in this chapter. Lastly, directives play an important role and are an important driver for SDI development – especially in the marine domain. Thus the most important directives are stated at the end of this chapter.

Chapter 3 – Existing approaches and established systems Using technology and standards of the preceding chapter systems and software tools were built that are described in this chapter. Firstly MDI-DE is introduced as this was developed together with this thesis and is a completed project now. After that other (international) MSDI approaches are presented. Furthermore existing reference models are described. Thesauri and controlled vocabulary are an aspect of SDIs and spatial data which is why existing vocabularies and tools to convert and present vocabularies on the web are depicted. The chapter closes with an overview over tools to evaluate performance and conformity of services because services are an integral part of SDIs.

Chapter 4 – Evaluation of existing MSDIs Firstly this chapter builds a framework that enables (to a certain degree) objective evaluation of SDIs. Evaluation of other existing SDIs (in contrast to MDI-DE which is built from scratch) is important because this makes them comparable and highlights potential pitfalls as well as aspects and concepts worth incorporating into the own approach. These pros and cons are elaborated in the following section that evaluates existing MSDIs.

Chapter 5 – Selected implementation aspects of an interoperable architecture This chapter represents the synthesis of the findings so far that are used as bases for further implementations of selected aspects. Firstly it communicates the lessons that can be learned from other (M)SDIs and builds a reference model for MDI-DE based on these findings. Because the reference model lists all the actors of MDI-DE which have data sets and more importantly services available an overview was needed at the beginning of the project. This overview was achieved through web forms and tables. After the existing services are known and after additional ones were set up based on existing data sets performance and conformity of the services plays an important role which is why the services are evaluated afterwards. The service evaluation showed that there is the need to visualize results of the Service Status Checker to simplify evaluation of services with it. The prerequisite to set up services are data and according metadata sets. Especially for metadata annotation thesauri are important but also for services (e.g. keywords) and the MDI-DE portal itself (e.g. search function). The next-to-last section formulates requirements to build a marine thesaurus that supports the functionalities just mentioned. These requirements form the base for the actual implementation of a marine thesaurus which is the last implementation of this chapter.

Chapter 6 – Future prospects The last chapter provides an outlook on what MSDIs will look like in the future. It also details what additional features can be implemented and how certain aspects of this thesis can be improved in the future.

Appendices – Selected listings & in-depth evaluations of MSDIs Appendix A provides listings of the implementations *SSCVisualizer* and *JSKOSify* as well as the forms and tables to analyse the

existing data sets and services. Appendix B documents the lengthy evaluations of the MSDIs of Australia, Canada, the UK and the USA as well as a sort of self-assessment with Germany.

2 Fundamentals, basic concepts and standards

> *"Interoperability among components of large-scale, distributed systems is the ability to exchange services and data with one another."* (Heiler, 1995, p. 271)

Interoperability is the base for SDI development, i.e. makes it possible. (Staub, 2009, p. 20) states that interoperability has technical as well as organizational aspects which are depicted in figure 2.1. The figure also shows five elements that characterize interoperability and that make up the foundations for this thesis and that will be discussed in this chapter:

[1] Directives and laws
[2] Standards and norms
[3] Profiles/Data modelling
[4] Data transfer/Services
[5] Semantic transformation

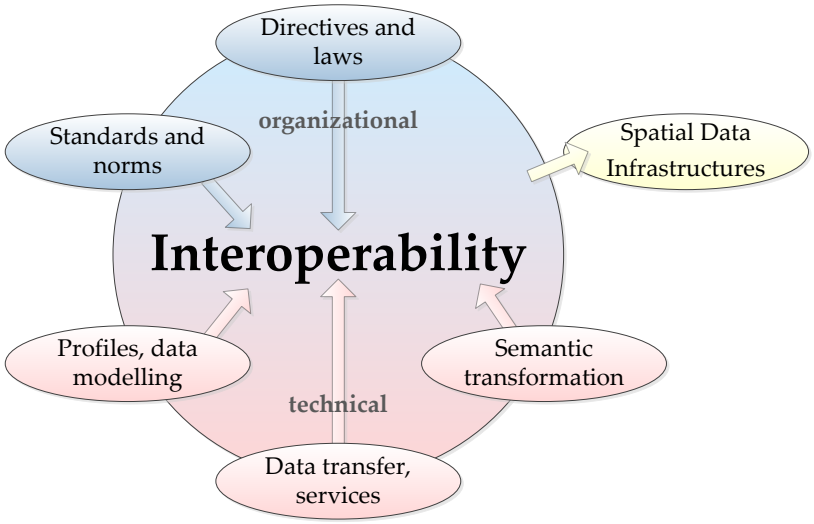


Figure 2.1: *Interoperability (modified after (Staub, 2009, p. 20))*

Chapter 1 already expressed the goal that several infrastructure nodes will be set up for the project which means that the participants retain control over and responsibility for their own data sets. An infrastructure with spatial data and a network of distributed nodes is called a spatial data infrastructure (SDI) and if the data sets handle spatial information in the marine domain it is called a marine (spatial) data infrastructure (M[S]DI). Both terms are explained in-depth in section 2.1. All these nodes use web services (*[4] data transfer/services*) so that the data owners do not have to push data files back and forth trying to keep track of which the current and most up-to-date version is. To make a SDI work the web services have to be able to communicate with each other. To achieve this web services have to be based on standards (*[2] standards and norms*) and these standards in the spatial domain are explained in section 2.2. Now that we have web services, standards for web services and nodes which are relying on web services, we must consider the architecture to compose the network of nodes. Because of all these services, it is certain that it will be or is a service-oriented architecture (also defined in section 2.2). This can be modelled with the help of a reference model based on the ISO standard Reference Model of Open Distributed Processing (RM-ODP) which is explained in section 2.3.

There are several European directives (*[1] directives and laws*) in the marine and SDI domain which affect many of the participants of the infrastructures mainly through reports with iterations over specified time spans and thus these directives will be discussed in detail in section 2.5.

So far data was referenced indirectly in almost every aspect addressed. But with data comes – or should come – metadata which describe the data and make it discoverable. But what if somebody wrote "caost" instead of "coast" as keyword for a metadata set? Nobody would be able to find this dataset by the keyword "coast". But maybe somebody who is interested in coastal data is also interested in data about beaches but cannot find it when searching with the keyword "coast". So what is needed besides data and metadata is knowledge representation (*[3] profiles/data modelling* and *[5] semantic transformation*) assuring that only words from a keyword list can be picked and connecting terms with other terms with a similar meaning. Important standards and approaches such as SKOS, RDF and ontologies are discussed in section 2.4.

2.1 Spatial Data Infrastructures

Spatial data infrastructures (SDIs) are about separated systems/nodes working together and interoperability is the ability of systems and/or organizations to work together. This means that interoperability is a prerequisite for SDI development which is why interoperability is discussed firstly (subsection 2.1.1). After that the term spatial data infrastructure will be defined (subsection 2.1.2) and from this definition components will be derived (subsection 2.1.3). This section closes with a classification of SDIs showing the broad fields a SDI can be applied to (subsection 2.1.4) with marine SDIs being picked as an example for a thematic SDI (subsection 2.1.5).

2.1.1 Interoperability

As already mentioned in the introduction of this chapter it is still a problem for many data owners that on the one hand they have files in various (spatial) formats causing incompatibility issues and on the other hand the files are stored on the workstations of the employees so that it is hard to keep track of the most up-to-date version of the file or the whereabouts of files and data in general. That is why interoperability is needed which is defined by (Heiler, 1995, p. 271) as

> *"[. . .] the ability [of systems] to exchange services and data with one another. It is based on agreements between requesters and providers on, for example, message passing protocols, procedure names, error codes, and argument types."*

As stated in the introduction of this chapter interoperability is divided into the two main areas *organizational* and *technical* interoperability[1]. Within technical interoperability the two characteristics *semantic* and *syntactic* interoperability are found. (Najar, 2006, p. 61) states that

> *"semantic Interoperability is a special kind of interoperability which provides systems with the ability of access, consistently and coherently, to similar (though autonomously defined and managed) classes of digital data, objects and services distributed across heterogeneous repositories."*

Since this definition is rather complex and specific the definition of (Kresse et al., 2012, p. 407) will be stated, too, which is more general and focuses on the user:

> *"Semantic interoperability is defined as the ability of a user to fully understand the data received in a data exchange in order to be able to make full use of those data if needed."*

(Danko, 2008, p. 657) simplifies semantic interoperability even more by stating that this is about

> *"[. . .] understanding the same term for the same concept."*

[1] cf. (Staub, 2009, p. 20)

This is the definition that will be used for the rest of this thesis (e.g. in section 2.4).

As the second technical characteristic of interoperability syntactic interoperability

> *"[. . .] allows the interoperable use of available data through a standard interface (through OGC web services). This interface is accessed through a standardized protocol and returns the information in a standard format. Query and delivery of data occurs in the structure of the provider model. The data structure of available data cannot be influenced by the user."*[2]

Standards are the base of a SDI which (Tóth et al., 2012, p. 20 and p. 22) underlines by stating:

> *"Interoperability arrangements and data harmonisation in SDIs aim to eliminate incompatibility and inconsistency of data, thereby exempting the users from having to undertake onerous data manipulations before they start using data in their applications.*
> *[. . .]*
> *The interoperability in an SDI means that users are able to integrate spatial data from disparate sources "without repetitive manual intervention", i.e. the datasets they retrieve from the infrastructure follow a common structure and shared semantics."*

This emphasises that both the problems of technical (not thematic) incompatibility and inconsistency are ruled out by an SDI which is based on interoperable web services which are explained in section 2.2.

2.1.2 SDI definition

To define what a spatial data infrastructure is we already got some initial points from the citations of (Tóth et al., 2012) in subsection 2.1.1. In addition to that (Tóth et al., 2012, p. 21) states:

> *"SDIs should encompass the common spatial aspects constituting a generic location context for a wide variety of applications."*

Therefore we end up with three basic items to get to a definition of the term *spatial data infrastructure* which are all about the access and sharing an SDI incorporates:

- No data manipulations needed by users
- Datasets follow a common structure and shared semantics and are retrievable through the infrastructure
- SDIs should include common spatial aspects to offer a generic location context for a wide variety of applications

This, however, is just a very small excerpt to understand what an SDI is and what constitutes it. (McGranaghan, 2003) citing (Groot & McLaughlin, 2000) gives a broader but also very concise definition of the term SDI:

> *"Geospatial Data Infrastructure encompasses the networked geospatial databases and data handling facilities, the complex of institutional, organizational, technological, human, and economic resources which interact with one another and underpin the design, implementation, and maintenance of mechanisms facilitating the sharing, access to, and responsible use of geospatial data at an affordable cost for a specific application domain or enterprise."*

We again see the access and sharing of data aspect but also a wide range of other aspects such as:

- **Technology:** networked geospatial databases and data handling facilities
- **Organisation/Policy:** complex of institutional, organizational, technological, human, and economic resources
- **Cost:** affordable cost
- **People/Users:** specific application domain or enterprise

[2] cf. (Staub, 2009, p. 25) (translated)

2.1.3 Components of an SDI

As can be seen, the aspects are already categorized which suggests the possible components of an SDI. For two of the aspects the GSDI Cookbook (Nebert, 2004, p. 8) is giving more in-depth information in its comprehensive definition of the term. Firstly it describes the beforehand mentioned users of an SDI more precisely as *"[. . .] users and providers within all levels of government, the commercial sector, the non-profit sector, academia and by citizens in general"* and secondly it specifies the organizational or policy aspect in the sense that an SDI *"[. . .] must also include the organisational agreements needed to coordinate and adminster it on a local, regional, national, and or transnational scale."*

The findings so far align with the components of an SDI found in (Rajabifard. & Williamson, 2001, pp. 4) and which can be seen in figure 2.2: people, data, access network, policy and standards. Everything but standards and access networks were mentioned more or less directly but since access and sharing of data needs standards as well as access networks these two components were already implicitly included in the definitions so far.

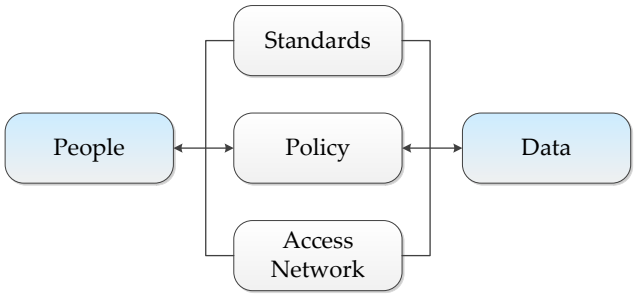


Figure 2.2: *Components of an SDI (modified after (Rajabifard. & Williamson, 2001))*

But the GSDI Cookbook (Nebert, 2004, p. 8) is digging deeper into the components of an SDI as it did with the other aspects mentioned. It lists the components as:

- **metadata** (geographic data and attributes, sufficient documentation),
- **catalogues and web mapping** (discovery, visualization, evaluation),
- access and
- additional services for data application.

This leads to an extended view of the components of an SDI and results in the refined figure 2.3.

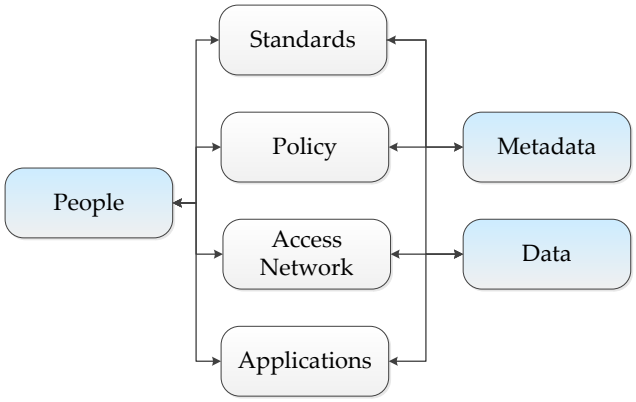


Figure 2.3: *Components of an SDI expanded with aspects of (Nebert, 2004) (modified after (Rajabifard. & Williamson, 2001))*

2.1.4 Classification of SDIs

(Rajabifard. & Williamson, 2001) also point out that there are SDIs at different political-administrative levels which make hierarchies the first form of classification. Figure 2.4 illustrates these levels and shows that there are vertical and horizontal relationships between the levels. (Rajabifard. & Williamson, 2001), however, only mention relationships on a horizontal level but do not show them in their figure. (Bernard et al., 2005, p. 7) is extending the original figure with these horizontal relationships.

The vertical relationships represent that a local SDI delivers data to a state SDI which is composed of many local SDIs. That is the part where the state SDI is facing down but it also has to face up because it has to deliver data to the SDI above it – the national SDI. In (Bill, 2010) examples for the different levels can be found making the levels and their relationships easier to understand:

- **global SDI** – Global Spatial Data Infrastructure
- **regional SDI** – INSPIRE
- **national SDI** – SDI for Germany (GDI-DE)
- **state SDI** – SDI for Mecklenburg-Vorpommern (GDI-MV)
- **local SDI** – GeoPortal of the city of Rostock

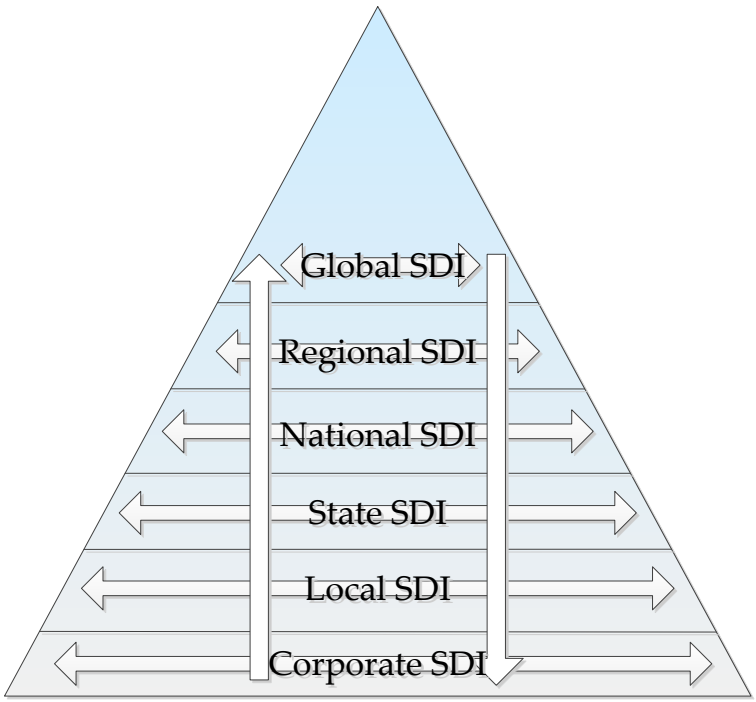


Figure 2.4: *SDI hierarchy with vertical and horizontal relationships (modified after (Rajabifard. & Williamson, 2001) and (Bernard et al., 2005))*

The hierarchical levels affect all of the components that were pointed out before. For instance when the national SDI is using a certain standard for its data or data access the state SDI below is very likely to use this standard as well. This is equally true for the horizontal relationships because it is more likely that two states will use the same standards rather than that a state is interested in using standards from a local SDI (Bernard et al., 2005, p. 7). In Europe (regional SDI) this is answered through the INSPIRE directive that specifies guidelines for the national SDIs (e.g. GDI-DE in Germany).

Another approach to classify SDIs is to look at their thematic scope. While there are many SDIs for specific data coverages such as urban planning respectively sustainable land management (Groot, 1997) or archaeological and built heritage (McKeague et al., 2012) which do not have specific identification respectively a name of their own we find the term *Environmental Spatial Data*

Infrastructure for example in (Fabian, 2003). Although it has to be stated that this term does not seem to be widely used for SDIs handling environmental data. In Germany, for instance, there is the PortalU which is an SDI for environmental data. However, in the marine domain a term evolved to describe these SDIs – *Marine Data Infrastructures* (MDI) or *Marine Spatial Data Infrastructures* (MSDI) or *Marine Geospatial Data Infrastructures* (MGDI, used in Canada). In the domain of Integrated Coastal Zone Management (ICZM) the term Coastal Spatial Data Infrastructure (CSDI) is often used, too. The widespread use of these terms can be seen in section 3.2 where different approaches on implementing MSDIs in many countries including Australia, Canada and Ireland are presented.

2.1.5 Marine SDIs

The term MSDI dates back to at least 2001 (Vessie et al., 2001) but most probably was used long before with the CoastGIS conference series' beginning in 1995. According to (Strain, 2006), MSDIs are about the exchange and sharing of spatial data like SDIs with the significant difference that SDIs are primarily focused on land-related data, while MSDIs are aiming at improved access to marine themed data to advance marine and coastal zone administration and management. Figure 2.5 shows some of the activities marine and coastal zone administration involves and which an MSDI has to cover.

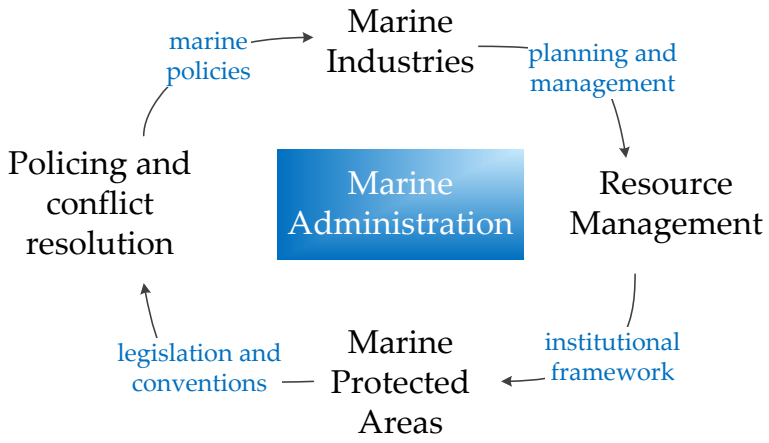


Figure 2.5: *Marine Administration (modified after (Strain et al., 2006))*

(Russell, 2009) is giving a quite comprehensive definition of the term MSDI with stating that an MSDI is

> *"[. . .] the component of a National SDI that encompasses marine and coastal geographic and business information in its widest sense. An MSDI would typically include information on seabed bathymetry (elevation), geology, infrastructure (e.g. wrecks, offshore installations, pipelines, cables); administrative and legal boundaries, areas of conservation and marine habitats and oceanography."*

(Bartlett et al., 2004, p. 6) is also arguing that it *"[. . .] is not possible to develop a coastal SDI in isolation from the broader national or regional SDI (NSDI)"* and that a *"[. . .] CSDI will typically be a subset of a more comprehensive NSDI because the coastal zone covers multiple physical and institutional spaces included in the generic NSDI."*

But it has to be indicated that an MSDI is not in all cases a component of a National SDI because (Strain, 2006) is also stating examples for MSDIs on a regional and global level. While not mentioning a coherent example for a regional MSDI she lists two global MSDI initiatives: Global Oceans Observing System (GOOS) and Oceans 21. An example for a regional MSDI (although not calling itself SDI or MSDI) is the Oregon Coastal Atlas[3] for instance.

[3] http://www.coastalatlas.net/

Now with the classification into the hierarchical system (global, national, regional) the other aspect outlined in (Rajabifard. & Williamson, 2001) – the components of a SDI (data, standards, policies, access networks and people) – has to be examined and checked for its applicability to the marine environment.

Generic **standards** for services like the ones by the OGC of course also apply to the marine domain. However, because the ISO TC/211 (see subsection 2.2.1 on page 24) is mostly focused on terrestrial spatial data standards for marine (meta)data are needed. Coordination is important to build standards. This is why the Intergovernmental Oceanographic Commission (IOC) was established in 1960 because[4] *"its mission is to promote international cooperation and to coordinate programmes in research, services and capacity building to learn more about the nature and resources of the oceans and coastal areas, and to apply this knowledge to improved management, sustainable development and protection of the marine environment and the decision making processes of States."* In order to achieve this the IOC established the International Oceanographic Data and Information Exchange (IODE) in 1961. IODE facilitates[5] *"[. . .] the exploitation, development, and exchange of oceanographic data and information between participating Member States and by meeting the needs of users for data and information products."* While the International Hydrographic Organisation (IHO) and the International Hydrographic Bureau developed a standard for hydrographic data (S-57). (Strain et al., 2006) approached the issue of standards for MSDIs, too, and state that the standard *marineXML* has been developed by the International Oceans Commission (IOC). However, this effort seems to be discontinued but may be still in use in Australia only. Furthermore there is the Hydrologic Markup Language (HydroML) which allows *"[. . .] the definition of hydrologic information"*[6] and XHdyro which *"[. . .] is an XML format for inter-departmental and cost-efficient time-series data exchange"*[7] developed by the German Federal Institute of Hydrology (BfG).

(Strain et al., 2006) also state that **Policies** are covering access, data custodianship, conformity, quality, content, industry engagement, avoidance of duplication and sensitivity. Except for data quality, data access and privacy all these fields are the same as for terrestrial (land-based) spatial data when applied to the marine domain. Data quality may be more difficult to achieve due to the complexity of the marine environment (complex measurements and processes). While data access is no problem onshore because of fixed line data transfer for offshore usage wireless data transfer may be needed which could be problematic. Because countries are reluctant to share spatial information relating to their marine jurisdictions different privacy policies for offshore data may be needed.

The issue of offshore data is also the only difference when it comes to **access networks** comparing terrestrial and marine data because the technology that is used for data transfer and access on land is not appropriate for offshore use. Examples for access networks in the marine domain include inter alia the Global Ocean Observing System (GOOS) managed by the IOC which provides[8] *"[. . .] a coordinated approach to deployment of observation technologies, rapid and universal dissemination of data flows and delivery of marine information to inform and aid marine management and decision makers and to increase the appreciation of the general public of our changeable oceans."* Further examples are the Integrated Ocean Observing System (IOOS) and the Global Monitoring for Environment and Security (GMES, now Copernicus[9]) initiative.

The importance of **people** in the marine domain is just like it is in the terrestrial domain. *"The key to success in SDI initiatives are partnerships within and between organisations involved in marine administration and spatial information."*(Strain et al., 2006)

(Strain et al., 2006) also state that **data** collection and updating is more difficult in the marine environment because it is dynamic and multi-dimensional (to a greater extent than land-based spatial data). It is also pointed out that there are two key issues when it comes to data which are the same

[4] http://ioc-unesco.org/index.php?option=com_content&view=article&id=14:about-the-ioc
[5] http://www.iode.org/index.php?option=com_content&view=article&id=385&Itemid=34
[6] http://water.usgs.gov/XML/NWIS/nwis_hml.htm
[7] http://www.xhydro.de/index_en.html
[8] https://en.unesco.org/node/119895
[9] www.copernicus.eu

that apply to SDIs: availability and interoperability. The source is also listing "fundamental datasets" for MSDIs:

- cadastral
- address
- transport
- administrative & political boundaries
- elevation
- hydrography
- imagery
- bathymetry
- marine protected areas
- oceanography
- sea level
- waves
- water quality
- sea floor composition
- meteorological conditions
- biodiversity regionalization

To see what further data sets are of interest the INSPIRE directive (in-depth description in subsection 2.5.1 on page 42) is used. (Korduan, 2013) analyses the coverage of marine data within the INSPIRE directive. He states that there are 19 themes important for the marine domain of which the most important ones are:

(1) Oceanographic Geographical Features (OF, e.g. sea surface temperature, currents, wave heights or salinity) (Millard et al., 2013a)
(2) Land Use (LU, use and functions of a territory, e.g. `1_4_AquacultureAndFishing`) (Salgé et al., 2013)
(3) Energy Resources (ER, offshore wind parks, energy derived from tidal movement, wave motion or ocean current) (Tuchyna et al., 2013)
(4) Mineral Resources (MR, mineral resources in or on the sea floor) (Serrano et al., 2013)
(5) Natural risk zones (NZ, marine related hazard types like floods) (Harrison et al., 2013)
(6) Environmental monitoring Facilities (EF, Oceanographic Geographical Features are derived from Environmental monitoring Facilities) (Daffner et al., 2013)
(7) Habitats and biotopes (HB, includes fresh water and marine areas) (Hinterlang et al., 2013)
(8) Bio-geographical regions (BR, "Areas of relatively homogeneous ecological conditions with common characteristics", e.g. Baltic sea) (Roscher et al., 2013)
(9) Sea Regions (SR, "A Sea Region is a defined area of common (physical) characteristics", e.g. coastline) (Millard et al., 2013b)
(10) Area management/restriction/regulation zones and reporting units (AM, "areas managed, regulated or used for reporting") (Lihteneger et al., 2013)
(11) Agricultural and Aquaculture Facilities (AF, e.g. marine and freshwater aquaculture) (Busznyák et al., 2013)

Figure 2.6 shows that some of the themes have relationships with each other and/or other non-marine specific themes. The connections are:

- Oceanographic Geographical Features
  → Oceanographic Geographical Features are derived from *Environmental monitoring Facilities* (EF)
  → Oceanographic Geographic Features always contain information about a *Sea Region* SR
- Sea Regions
  → Elevation (EL, depth of a Sea Region, not included in the eleven themes of (Korduan, 2013))
  → Main Sea Region class (`SeaArea`) derives from *Hydrography* (HY)
  → *Geographic Names* (GN) are used for the named Sea Regions
  → Geophysical observations (described by the *Oceanographic Geographical Features* [OF] theme) are made within Sea Regions
  → Areas of the sea may be *Area Management or Reporting Units* (AM)
- Area Management or Reporting Units
  → Areas of the sea (*Sea Regions* [SR]) may be Area Management or Reporting Units

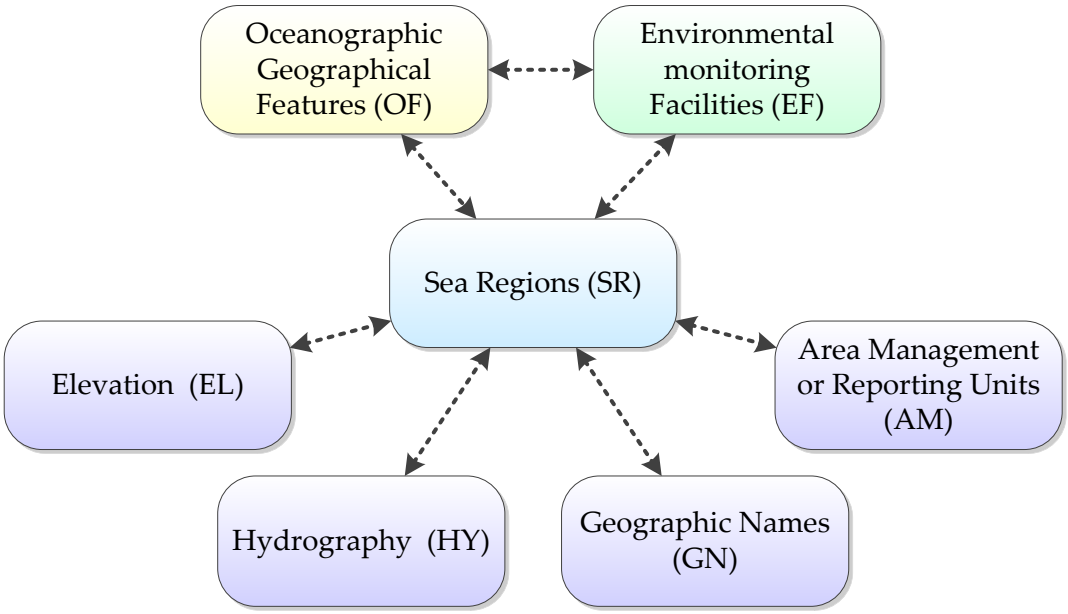


Figure 2.6: *Links between selected INSPIRE themes*

Except for imagery all the "fundamental datasets" by (Strain et al., 2006) can be mapped to INSPIRE themes relevant to the marine domain (Korduan, 2013). However, because not all of these data sets are specific to the marine domain they are grouped accordingly and are then mapped to INSPIRE themes (use table 2.1 for reference):

- not marine-specific
  - cadastral ⇄ LU
  - address ⇄ AD
  - transport ⇄ TN
  - administrative and political boundaries ⇄ SR
  - elevation ⇄ SR/EL
  - imagery
  - meteorological conditions ⇄ NZ
  - biodiversity regionalization ⇄ HB and BR
- marine-specific
  - bathymetry ⇄ SR/EL
  - hydrography ⇄ SR/HY
  - marine protected areas ⇄ HB
  - oceanography ⇄ OF
  - sea level ⇄ OF
  - waves ⇄ OF
  - water quality ⇄ OF
  - sea floor composition ⇄ OF

Concluding this shows that the two main differences of the data components between marine and terrestrial environments are fundamental (marine-specific) datasets and the data collection process. Furthermore scientific data plays a much more prominent role in marine SDIs than it plays in terrestrial SDIs.

Table 2.1: *Selected INSPIRE themes and their abbreviations (INSPIRE, 2007)*

AF – Agricultural and Aquaculture Facilities	AD – Adresses
AM – Area Management or Reporting Units	EL – Elevation
OF – Oceanographic Geographical Features	LU – Land Use
EF – Environmental monitoring Facilities	SR – Sea Regions
BR – Bio-geographical regions	HY – Hydrography
HB – Habitats and biotopes	ER – Energy Resources
TN – Transport Networks	NZ – Natural risk zones
GN – Geographic Names	MR – Mineral Resources

2.2 Geospatial standards

SDIs (see section 2.1) rely on standards because they build on web services which – in the SDI world – were specified by the Open Geospatial Consortium (OGC, with standards such as WMS and WFS, see subsection 2.2.3) in conjunction with the efforts of the International Organization for Standardization (ISO, technical committee 211, see subsection 2.2.1). Data and services require metadata to be retrievable and easily accessible. For this reason subsection 2.2.2 examines metadata from its roots to internationally accepted standards.

2.2.1 ISO TC 211 and its 191XX series

National efforts like the Content Standard for Digital Geospatial Metadata (CSDGM) were on the one hand superseded by and on the other hand incorporated into an international agreement on geospatial metadata standards. In 1994 the ISO formed a technical committee (TC 211) to develop such an international agreement standardizing information related to the spatial domain. ISO/TC 211 is responsible – according to their overview website[10] – for

> *"[. . .] standardization in the field of digital geographic information [which] aims to establish a structured set of standards for information concerning objects or phenomena that are directly or indirectly associated with a location relative to the Earth."*

The outcomes of its work are the 191XX series of international standards of which selected ones with relevance to this thesis will be explained further in the next few paragraphs and ISO 19115 will be discussed in subsection 2.2.2.

19119 – Services play an important role in the world of SDIs because SDIs are collections of distributed services. With ISO 19119 the ISO wanted to standardize services by[11]

(1) providing an abstract framework to allow coordinated development of specific services,
(2) enabling interoperable data services through interface standardization,
(3) supporting development of a service catalogue through the definition of service metadata,
(4) allowing separation of data instances & service instances,
(5) enabling use of one provider's service on another provider's data and
(6) defining an abstract framework which can be implemented in multiple ways.

(3) is of special importance for the development of an SDI because in general metadata catalogues assist users searching for spatial data. With service metadata users can search what data a service offers. But ISO 19119 is also specifying metadata about services and not just data enabling users to find services as well. The metadata of services document among others states which requests the service supports, which layers it offers and what coordinate reference systems are used (Müller

[10] http://www.isotc211.org/Outreach/Overview/Overview.htm
[11] cf. (ISO, 2001, p. 4)

et al., 2004, p. 126). Subsection 2.2.3 will outline how the OGC built upon this[12] standard and developed specifications for the implementation of services like web map service (WMS) and web feature service (WFS).

19156 – Observations and measurements Especially in the marine domain much data originates from sensors which is why ISO 19156 – observations and measurements (O&M) – is of major importance to MSDIs. Although being an ISO standard the OGC was involved in developing O&M implementation specifications. On their website the ISO characterizes O&M in this way[13]:

> *"ISO 19156:2011 defines a conceptual schema for observations, and for features involved in sampling when making observations. These provide models for the exchange of information describing observation acts and their results, both within and between different scientific and technical communities."*

The term *conceptual schema* is defined by (Castano et al., 1998, p. 290) as a composition of elements and links. Whereat

> *"an element abstracts the constructs used in conceptual models to describe classes of real-world objects (e.g., entity, class). A link abstracts the constructs used in conceptual models to describe relationships between real-world objects due to the aggregation and generalization abstraction mechanisms (e.g., relationships, "is-a" links)."*

19136 – Geography Markup Language (GML) So far the ISO standards defined metadata and conceptual schemes how to store data. However, with GML the ISO also offers a standard on how to store spatial objects. GML was originally developed by the Open Geospatial Consortium (OGC, see subsection 2.2.3). Because it got widely adopted and used the GML specification has been incorporated into ISO's range of international standards concering spatial data (191XX series). The Encyclopedia of GIS (Raimundo & Chang-Tien, 2008) defines GML as follows:

> *"Geography Markup Language (GML) is an open-source encoding based on the eXtensible Markup Language (XML), and suitable for the representation of geographical objects. Organized as a hierarchy of features, collections, and geometries, among other structures, GML objects are modeled after real-world entities characterized by properties and state."*

Furthermore GML is used as an information exchange and storage format for data sharing by defining a schema of how spatial data can be characterized so that systems are able to understand each other. This schema is the framework for the data and has to be distinguished from the actual data which is the case with most XML applications (Raimundo & Chang-Tien, 2008, p. 364).

2.2.2 Metadata standards

Without additional data a river for instance would just be some line geometry and it could not be differentiated from a street. Only with metadata (and categorization and attribution) one can differentiate the two different lines. Just having metadata at all is great when staying in the realm of isolated systems but when it comes to interoperability and systems interacting with each other a standard is needed which defines what metadata have to look like.

Predating geospatial data librarians were the first using computers to catalogue their data (i.e. books and other physical media). For interoperability the machine-readable cataloging (MARC) standard evolved. Due to its complexity and largeness the *Dublin Metadata Core Element Set* (or Dublin Core, for short) was developed in March 1995 which only has 13 data elements. According to (Guptill, 1999, p. 682) it

[12] And of course on other ISO standards like 19136 (GML) as well.
[13] http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=32574&commid=54904

> *"[. . .] was proposed as the minimum number of metadata elements required to facilitate the discovery of document-like objects in a networked environment such as the Internet."*

However, even before Dublin Core was developed the US Federal Geographic Data Committee (FGDC) proposed the *Content Standard for Digital Geospatial Metadata* (CSDGM) in June 1994. (Guptill, 1999, p. 683) states that

> *"The standard was the first focused effort on specifying the information content of metadata for a set of geospatial data. The standard was developed from the perspective of defining the information required by a prospective user to determine the availability of a set of geospatial data, to determine the fitness of a set of geospatial data for an intended use, to determine the means of accessing the set of geospatial data, and to transfer successfully the set of geospatial data."*

A rather new standard for geospatial metadata succeeding the so far mentioned standards is *ISO 19115* which[14]

> *"[. . .] defines the schema required for describing geographic information and services. It provides information about the identification, the extent, the quality, the spatial and temporal schema, spatial reference, and distribution of digital geographic data."*

With its over 400 metadata elements – arranged in packages such as *reference system information*, *metadata extension information*, *data quality information* and *content information* and which can be mandatory, conditional or optional – ISO 19115 is enabling interoperability. *ISO 19139* is built upon these definitions and defines an XML Schema implementation for them (Bartelme, 2005, pp. 380).

2.2.3 OGC specifications

The *Open Geospatial Consortium* (OGC) was founded in 1994 and was known as the *OpenGIS Consortium* until 2004. It is an consortium of economic, government and research-based organizations which mission is to advance the development and use of GIS and spatial data by creating common, open standards and specifications that enable interoperability (Lupp, 2008, p. 815). In addition to the formal specification languages such as the Geography Markup Language (GML, see ISO 19136 on page 25) the results of the work of the OGC are primarily *OpenGIS Implementation Specifications* which define open interfaces and protocols. Products that conform to these specifications ensure interoperability.

In the following, specifically two of the OGC Web Services – abbreviated *OWS* – are discussed namely the Web Map Service (WMS) and Web Feature Service (WPS). Other important specifications include the Web Processing Service (WPS) that executes processes (such as buffer, overlay etc.) with input from another service (such as WFS) as well as the Catalog Service for the Web (CSW) that provides geospatial metadata and search thereof.

A **Web Map Service (WMS)** returns maps of spatially referenced data. The maps are produced dynamically from geographic information. A ***map*** is

> *"[. . .] a portrayal of geographic information as a digital image file suitable for display on a computer screen. A map is not the data itself."*[15]

Maps are typically returned to the user in a raster data format such as PNG, GIF or JPEG but can also be delivered in a vector-based format such as Scalable Vector Graphics (SVG) or Web Computer Graphics Metafile (WebCGM). A WMS is invoked by submitting a request with a special URL. What this URL looks like depends on the desired operation. There are three operations offered by a WMS of which one is optional:

- GetCapabilities: Service returns a XML document describing the service. Inter alia metadata about the service such as title, responsible party and so on as well as information about the

[14] (ISO, 2002b)
[15] (de la Beaujardiere, 2006, p. v)

offered layers (name, supported coordinate reference systems [CRS] etc.) are the output of this operation.
- GetMap: Service returns a map with the given geographic and dimensional parameters.
- GetFeatureInfo (optional): Service returns information about particular objects (features) from the map.

In case of a GetMap request the URL includes parameters indicating which area is mapped (`BBOX`), the width and height of the output image, the CRS and what data will be depicted on the map (`LAYERS`). This leads to a request such as this[16]

```
http://gdisrv.bsh.de/arcgis/services/CONTIS/Administration?
REQUEST=GetMap
&SERVICE=WMS
&VERSION=1.3.0
&CRS=CRS:84
&BBOX=3.0,53.0,20.0,55.0
&LAYERS=7,6,5,4,3,2,1
&WIDTH=640
&HEIGHT=400
&FORMAT=image/png
```

The response to this request is depicted in figure 2.7 and shows multiple layers. A ***layer*** is a

> *"basic unit of geographic information that may be requested as a map from a server"*[17]

A layer can also be defined as a set consisting of at least one feature. When two or more layers (or maps which can include multiple layers each) sharing the same geographic parameters and output size are combined an overlay can be produced (like the one shown in figure 2.7).

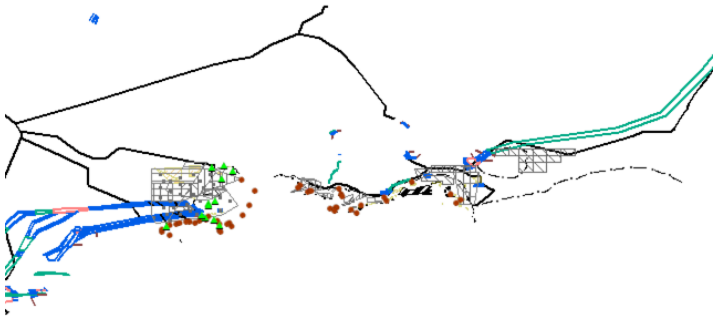


Figure 2.7: *Image response of a GetMap request to a WMS*

In contrast to a WMS a **Web Feature Service (WFS)** works with and outputs vector data, i.e. features. According to the WFS Implementation Specification[18] a ***feature*** is an *"abstraction of real world phenomena"*. A WFS is chosen over a WMS if geospatial operations will be performed on the data, for instance in order to create a buffer. A WFS goes beyond end user visualization but can be used for this as well (Michaelis & Ames, 2008, p. 1261). The four different WFS types (or *conformance classes*) are categorized according to the operations they support. The *simple WFS* (as well as the *basic WFS*[19]) implements the operations

[16] Taking the WMS "Continental Shelf Information System" by the Federal Maritime and Hydrographic Agency (BSH) as an example.
[17] (de la Beaujardiere, 2006, p. 7)
[18] (Vretanos, 2010, p. 4)
[19] The two only differ in the way the `GetFeature` opertation is executed.

- GetCapabilities,
- DescribeFeatureType,
- ListStoredQueries,
- DescribeStoredQueries and
- GetFeature.

The `GetCapabilities` operation is similar to the WMS. While `DescribeFeatureType` illustrates the structure of a particular feature `GetFeature` returns specific features in GML format, i.e. it returns geodata in vector form. The two other operations handle stored query expressions. According to (Vretanos, 2010, p. 30) a query expression

> *"[. . .] is an action that directs a server to search its data store for resources that satisfy some filter expression encoded within the query."*

Furthermore a stored query expression

> *"[. . .] is a persistent, parameterized, identifiable query expression. A stored query can be repeatedly invoked using its identifier with different values bound to its parameters each time."*[20]

The operation `ListStoredQueries` lists the available stored queries and `DescribeStoredQueries` returns detailed information about stored queries. In addition to these operations a *transactional WFS* also implements the `Transaction` operation. This operation enables *"[. . .] clients [to] create, modify, replace and delete features in the web feature service's data store."*[21] A *locking WFS* furthermore implements the operation/s `GetFeatureWithLock` and/or `LockFeature`. `LockFeature` is used to ensure serializability in transactions which means that a feature cannot have been altered by another user while it is modified by a user because the user firstly locks the feature. Afterwards the user can request a feature by using the `GetFeatureWithLock` operation and can safely modify it (Vretanos, 2010) and (Sinha, 2008).

When bringing the findings regarding SDI components (subsection 2.1.3 on page 18) and standards together a good overview develops. The components, how they interact and what standards are important for them are shown in figure 2.8.

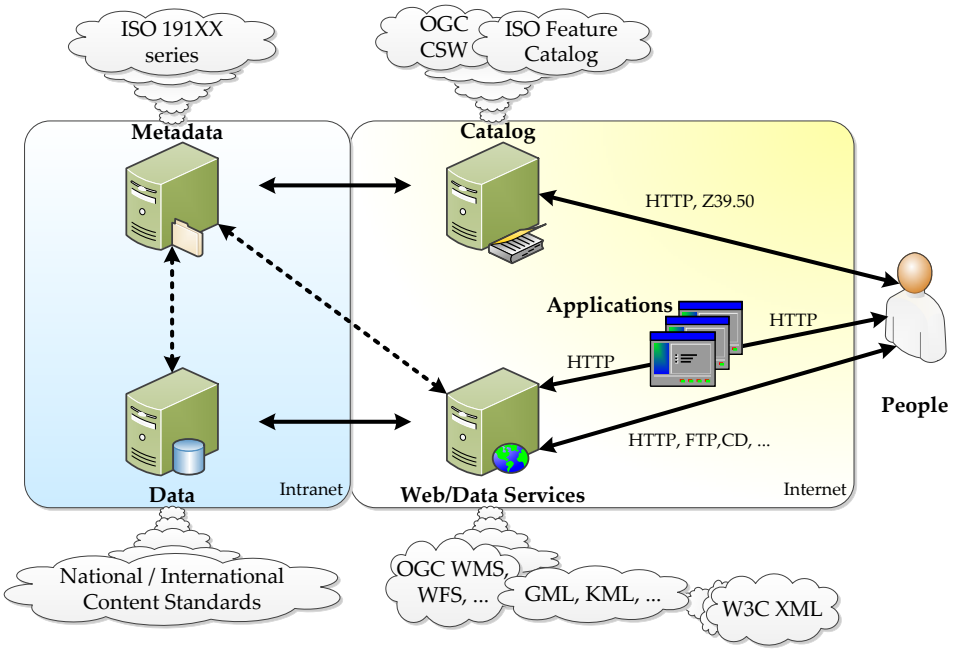


Figure 2.8: *Components, their interaction and standards in SDI (inspired by (Nebert & Anthony, 2010, p. 57))*

[20] (Vretanos, 2010, p. 42)
[21] (Vretanos, 2010, p. 90)

2.3 Standards for reference models

Section 2.2 described the standards which are the foundations of an SDI. The practical implementation of an SDI may be aided by a model. Because this model is giving the SDI a framework and because it is the foundation of it, it is called a reference model.

The are several reference models in existence that are out-dated, such as *Purdue Enterprise Reference Architecture* (PERA); *Process, Organization and Location and Data, Applications and Technology* (POLDAT) and the *Open-system environment* (OSE) *reference model* (RM). Other reference models fit only specific fields, e.g. business with models such as *Workflow Reference Model*, *Business reference model* and to some degree the *Business Process Execution Language* (BPEL) by the Organization for the Advancement of Structured Information Standards (OASIS). Another OASIS approach is the *SOA Reference Model*. SOA stands for service-oriented architecture. Because a SDI is a service-oriented architecture this reference model might be of interest. However, because it focuses solely on the architectural aspects and in particular services it does not fit the broad SDI development.

A standard for reference models that represents all aspects (with its viewpoints) of SDI development is the *Reference Model of Open Distributed Processing* (RM-ODP, see subsection 2.3.1) by the International Organization for Standardization (ISO). A similar earlier approach (*The "4+1" View Model of Software Architecture*) will be discussed in subsection 2.3.2. Because today modelling depends heavily on standardized notation to ensure interoperability the relationship between the Unified Modeling Language (UML) and the approaches for reference models is outlined in subsection 2.3.3.

2.3.1 RM-ODP

The purpose of building a reference model is to define a framework which structures large and complex distributed systems for which spatial data infrastructures are an example (Vallecillo, 2001, p. 2). The base for such efforts can be ISO's Reference Model of Open Distributed Processing (RM-ODP)[22]. Because it is a well-defined standard by the ISO RM-ODP was chosen as basis for implementing the interoperable infrastructure for the MSDI of Germany (MDI-DE). Furthermore according to (Hjelmager et al., 2008, p. 3) RM-ODP was already widely adopted as the conceptual base for other reference models like ISO standard 19101 (Geographic Information – Reference model (ISO, 2002a)), the OGC Reference Model which states in a deprecated version (Percivall, 2003, p. 3) that RM-ODP is applied in two ways: *"1) a way of thinking about architectural issues in terms of fundamental patterns or organizing principles, and 2) a set of guiding concepts and terminology."* and the Geospatial Interoperability Reference Model (G.I.R.M.) by the Federal Geographic Data Committee (FGDC) which is using the computational and information viewpoint of RM-ODP. Viewpoints are the core of RM-ODP enabling to focus on specific parts of an architecture or framework. Viewpoints are necessary because different stakeholders or actors have distinct interests in a system. While some aspects are relevant to developers they certainly are not relevant for customers. Taken classes as example – these are interesting for developers but customers are more interested in what the system provides them and not in the technical details such as classes (Staveley, 2011).

Viewpoints As depicted in figure 2.9[23] there are five generic and complementary viewpoints on the system to be modelled and its environment which (Vallecillo, 2001, p. 3) and (ISO, 2009, p. 5) describe as:

- **enterprise viewpoint** – *What for? Why? Who? When?*
  - focuses on the purpose, scope and policies for the system
  - describes the business requirements and how to meet them

[22] see (ISO, 1998b), (ISO, 1996a), (ISO, 1998a) and (ISO, 1996b)
[23] cf. (ISO, 2009, p. 5) and http://en.wikipedia.org/wiki/File:RM-ODP_viewpoints.jpg

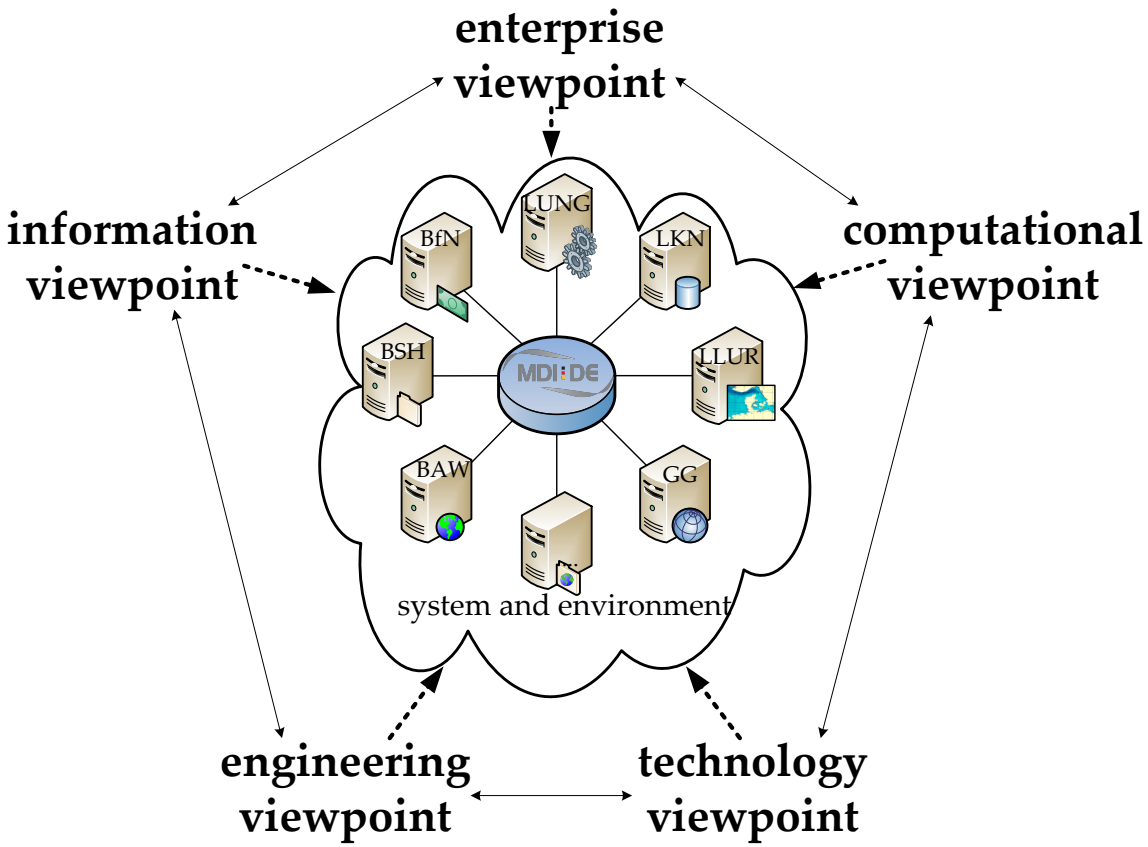


Figure 2.9: *RM-ODP's generic and complementary viewpoints on the system and its environment*

- **information viewpoint** – *What is it about?*
  - focuses on the semantics of the information and the information processing performed
  - describes the information managed by the system and the structure and content type of the supported data
- **computational viewpoint** – *How does each bit work?*
  - enables distribution through functional decomposition on the system into objects which interact at interfaces
  - describes the functionality provided by the system and its functional decomposition
- **engineering viewpoint** – *How do the bits work together?*
  - focuses on the mechanisms and functions required to support distributed interactions between objects in the system
  - describes the distribution of processing performed by the system to manage the information and provide the functionality
- **technology viewpoint** – *With what?*
  - focuses on the choice of technology of the system
  - describes the technologies chosen to provide the processing, functionality and presentation of information

2.3.2 The "4+1" View Model of Software Architecture

Another approach for describing an architecture which is also based on viewpoints was introduced by Philippe Kruchten in 1995 (Kruchten, 1995). The aim was the same as RM-ODP's – splitting the different aspects of a system into multiple views and describing an architecture with these allowing to address requirements of the different stakeholders. To achieve this goal he proposed these five (4+1) main views:

(1) **logical view**: object model of the design, contains information about the parts of the system
(2) **process view**: captures the concurrency and synchronization aspects of the design, encompasses some non-functional requirements such as performance and availability, too
(3) **physical view**: describes the mapping(s) of the software onto the hardware and reflects its distributed aspect i.e. by specifying the amount of nodes
(4) **development view**: describes the static organization of the software in its development environment, focusses on software modules and subsystems
(5) **use case view**: discovers the architectural elements and validates and illustrates the architecture

The use case view is the reason why this design approach is called *4+1* because the use case view is essentially redundant (hence *+1*). However, all the other views would not have been possible without it because the use cases respectively scenarios are an abstraction of the most important requirements leaving out specific details which means that the other views evolve on this base (Staveley, 2011).

2.3.3 Use of UML in reference models

Unified Modeling Language (UML) The aim of UML which is a standard specified by the Object Management Group

> *"[. . .] is to provide system architects, software engineers, and software developers with tools for analysis, design, and implementation of software-based systems as well as for modeling business and similar processes."*

The quote already indicates that UML has a very broad scope with many domains it may be applied to (OMG, 2011, p. 1). This is reflected by the variety of UML diagrams available. Diagrams give extensive information about a system in a graphical representation but in most cases this representation displays only part of the systems (a subset of its classes, components etc.) (OMG, 2011, p. 15).

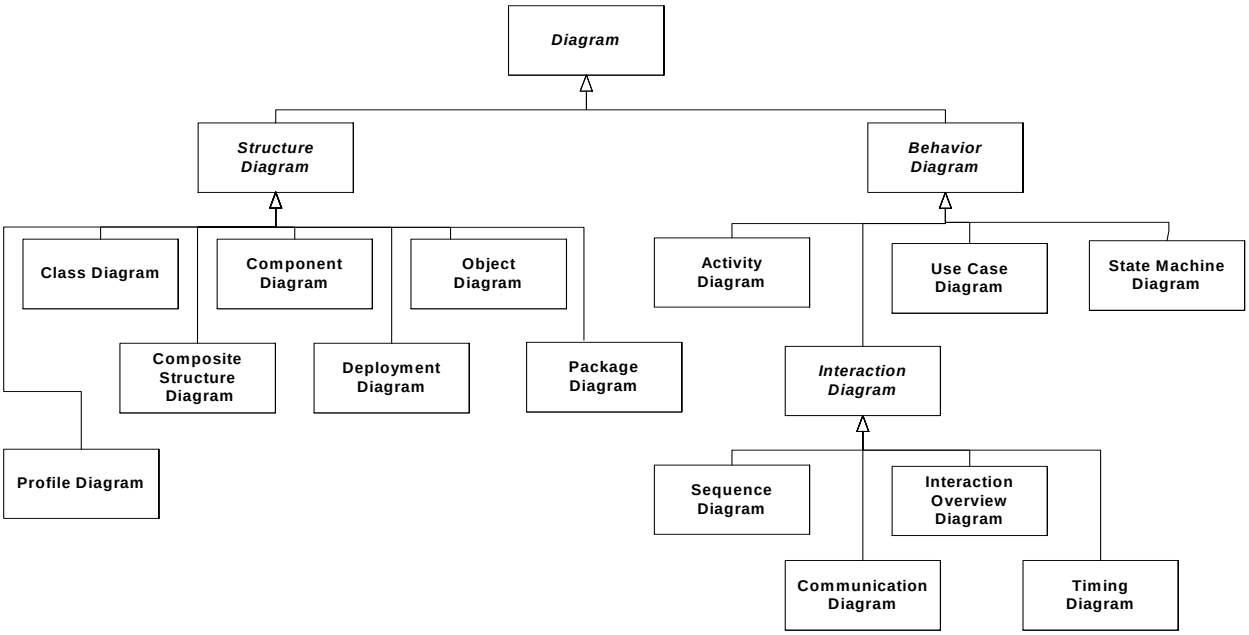


Figure 2.10: *Class diagram of UML diagram types (OMG, 2011, p. 694)*

As depicted in figure 2.10 the UML diagrams fall into two main categories: structure and behavior diagrams. In contrast to behavior diagrams which are dynamic, in the sense that they show interaction between elements, structure diagrams are static. Thus they only represent elements which are independent of time and which have to be available in the system being modelled. Taken class

diagrams as an example for structure diagrams they specify the classes, their attributes and the relationships between the classes of the system (OMG, 2011, p. 694).

As already stated behavior diagrams are dynamic thus pointing out how the system changes over time. Taken use case diagrams as an example for behavior diagrams they describe the functionality of a system in regard to actors, their goals which are represented as use cases and relationships and dependencies between these use cases (OMG, 2011, p. 694).

RM-ODP The original documents for the RM-ODP standard mention neither UML as a tool for modelling a system respectively an infrastructure nor a notation or model development method. But since UML gained importance in the last years the ISO proposed a standard for the *"Use of UML for ODP system specifications"* (ISO, 2009) at a later stage. The standard describes and defines how the viewpoints can be modelled with UML. For three of its five viewpoints it suggests diagram types which should be used to model certain aspects within these viewpoints. To model the configuration and structure of computational objects and their dependencies, composition and decomposition within the *computational viewpoint* the standard recommends a component diagram. Activity diagrams, state charts and interaction diagrams should be used to model interactions between computational objects (ISO, 2009, p. 44).

Within the *engineering viewpoint* a configuration of engineering objects which are structured as clusters, capsules or nodes is expressed by component diagrams, the activities going on within them with activity diagrams and the interactions between the engineering objects with sequence, activity and interaction diagrams. Component diagrams are also used to model the structure of a node (ISO, 2009, p. 58).

The *technology viewpoint* models how specifications are implemented using component diagrams. It also models the structure of node instances and communication links between them using deployment diagrams (ISO, 2009, p. 60).

The "4+1" View Model of Software Architecture While (Kruchten, 1995) proposed a notation for each viewpoint he could not have made proposals regarding UML diagrams to use for the viewpoints because the "4+1" View Model was put forward two years before UML was developed. However, inter alia (Staveley, 2011) and (Kontio, 2005) make proposals on which UML diagrams can be used for each viewpoint[24]:

- logical view: ***class***, *object*, *state machine*, *interaction* (e.g. ***sequence***) and **communication** diagrams
- process view: ***activity*** diagrams
- physical view: ***deployment*** diagrams
- development view: ***package*** and *component* diagrams
- use case view: ***use case*** diagrams

2.4 Standards for knowledge representation

The GSDI Cookbook (Nebert, 2004, p. 8) mentioned metadata as one important component of an SDI (see subsection 2.1.3 on page 18). Metadata is data about data – metadata describes the data by stating its owner, its thematic scope, how the data was collected, how often it is updated and so on. All that is needed is some form with some blank fields where the user types the information with which he likes to describe the data (set). In order to index the metadata fields *terms* are needed.

Semantic interoperability (see subsection 2.1.1 on page 16) is needed to ensure that different actors or systems have a common understanding of the meanings of terms. Interoperability was already defined by (Heiler, 1995, p. 271) (in subsection 2.2.1 on page 24) to be *"[. . .] the ability to exchange services and data with one another. It is based on agreements between requesters and providers on, for example,*

[24](Staveley, 2011) in *italics*, (Kontio, 2005) in **bold** and both in ***italics and bold***

message passing protocols, procedure names, error codes, and argument types." Semantic interoperability makes sure that both the requesters and providers have the same understanding of said services and data.

Let us assume a user is indexing metadata for a data set which is about a beach and there is a metadata field *keywords* where he accidentally puts "baech" instead of "beach". If this mistake is overlooked nobody who is looking for data about beaches will be able to find his data set because of the typo. If he would have had to choose from a list of keywords or by autocompleting his typing – i.e. by using a controlled vocabulary respectively a predefined list of terms – this problem would not have happened. Synonyms are another example to encourage the use of a controlled vocabulary respectively a thesaurus. When a thesaurus is used in conjunction with the search function the typo problem is eliminated as explained above. A thesaurus can also contain relationships between terms to cover synonyms for example. Continuing the "beach" example it would be great for a user that when he is looking for data about beaches he gets results which contain "coast", too, because the terms are used interchangeably at times, although scientifically they represent a hierarchy and cannot be used interchangeably. But a user with little knowledge might want to get data about "coast" although he/she is searching for the term "beach".

The superordinate concept in the field of knowledge representation is *ontology* which is explained in subsection 2.4.2. Formal languages describing ontologies are inter alia the Resource Description Framework (RDF, see 2.4.2.1 on page 35)[25] and the Web Ontology Language (OWL) which are specifications by the World Wide Web Consortium (W3C).

Neither RDF(S) nor OWL are intended or specialized for use with vocabularies and thus only offer limited labelling capabilities and – especially in case of OWL – have semantically strict relationships (super-/subclasses, not weaker ones like "related").

However, the Simple Knowledge Organisation System (SKOS, see 2.4.2.2 on page 37) is specified by the W3C to organize knowledge and model thesauri in RDF. The *simple* in its name means that it is very easy to map concept trees and relations. These facts make SKOS a good choice to represent controlled vocabularies on the web.

Besides SKOS there are also several other standards available to represent vocabularies that are also built on RDF. However, these are specified for other fields such as DOAP (Description of a Project), FOAF (Friend of a friend) and SIOC (Semantically-Interlinked Online Communities Project) and thus are not general enough.

2.4.1 Fundamentals – XML and DOM

The foundations for all the approaches in the field of knowledge representation presented in this thesis were laid by the World Wide Web Consortium (W3C) in the form of the Extensible Markup Language (XML) which it defines as[26]:

> *"[. . .] a simple, very flexible text format [. . .]. [. . .] XML is also playing an increasingly important role in the exchange of a wide variety of data on the Web and elsewhere."*

Furthermore the W3C states about the structure respectively components of XML in its technical report respectively specification (Bray et al., 2008) that:

> *"XML describes a class of data objects called XML documents [. . .]. XML documents are made up of storage units called entities [. . .]."*

To work with these documents the W3C developed the Document Object Model (DOM) which[27]:

> *"[. . .] is a platform- and language-neutral interface that will allow programs and scripts to dynamically access and update the content, structure and style of documents. The document can*

[25] To be precise RDF Schema (RDFS) is meant here because RDF is a whole family of W3C specifications.
[26] http://www.w3.org/XML/
[27] http://www.w3.org/DOM/

be further processed and the results of that processing can be incorporated back into the presented page."

Later on in this thesis DOM will be used to implement a tool which converts Excel lists respectively Comma-separated values (CSV) files to SKOS format (see subsection 3.5.2 on page 64). Because the tool is implemented in Java a Java implementation of the DOM Application Programming Interface (API) will be used – JDOM that is according to (Harold, 2002):

"[. . .] an open source, tree-based, pure Java API for parsing, creating, manipulating, and serializing XML documents."

2.4.2 Ontologies

As already pointed out, ontologies are the superordinate concept in the domain of knowledge representation and (Gruber, 2009) states that

"[. . .] an ontology defines a set of representational primitives with which to model a domain of knowledge or discourse. The representational primitives are typically classes (or sets), attributes (or properties), and relationships (or relations among class members). "

This definition already mentions *domain of knowledge* and *representational primitives* (inter alia classes and relationships). (Jepsen, 2009) adds to this by specifying

"[. . .] that an ontology is a method of representing items of knowledge (ideas, facts, things – whatever) in a way that defines the relationships and classifications of concepts within a specified domain of knowledge."

While this definition also mentions the most important elements from the previous definition (domain of knowledge, concepts [classes] and relationships), it adds one further element: *items of knowledge*.

Components of and an example for an ontology The two definitions delivered a number of components such as concepts/classes, items of knowledge and relationships. (Lord, 2010) states that

"Concepts, also called Classes, Types or Universals are a core component of most ontologies. A Concept represents a group of different Individuals, that share common characteristics, which may be more or less specific."

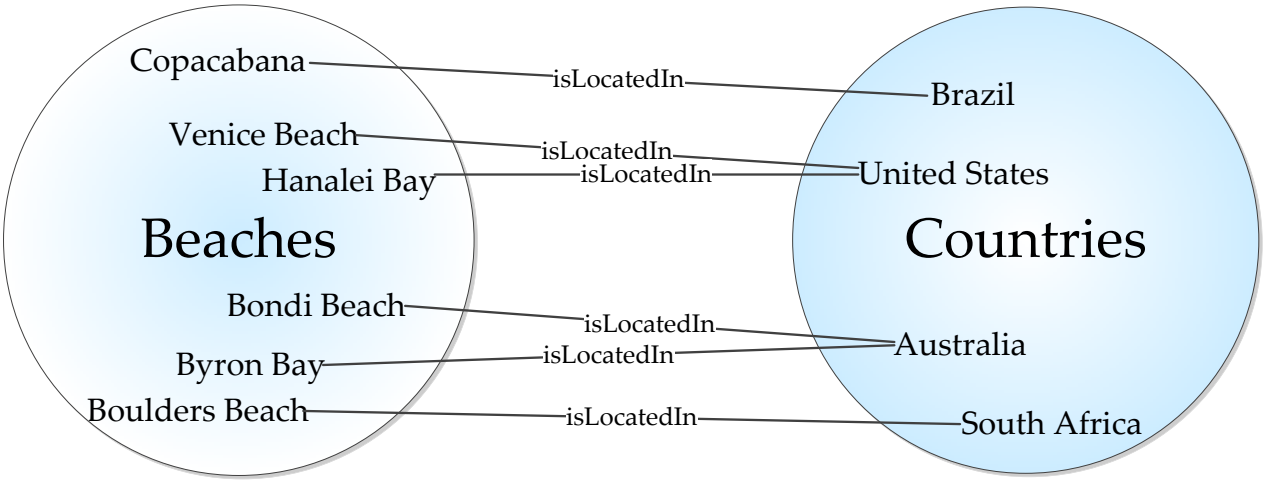


Figure 2.11: *Beaches and countries example illustrating relationships among classes and instances (inspired by (Jepsen, 2009))*

The example depicted in figure 2.11 is about beaches located in different countries. Beaches and countries are classes in this example. These two classes of course represent a set of individuals – a set of beaches such as Boulders Beach, Venice Beach, Hanalei Bay, Copacabana, Bondi Beach and Byron Bay as well as a set of countries such as Australia, Brazil, the United States and South Africa. Individuals are defined in (Lord, 2010) as

> *"[. . .] instances or particulars are the base unit of an ontology; they are the things that the ontology describes or potentially could describe. Individuals may model concrete objects such as people, machines or proteins; they may also model more abstract objects such as this article, a person's job or a function."*

When we stick to our example the relationship between the two classes is pretty easy to see: a beach has to be located in one of the countries (relationship *isLocatedIn* in figure 2.11). Again, (Lord, 2010) is offering a definition for the term relationships by pointing out that they

> *"[. . .] describe the way in which individuals relate to each other. Relations can normally be expressed directly between individuals (this article has author Phillip Lord) or between Concepts (an article has author a person) [. . .]."*

(Gruber, 2009) is taking us back to the W3C with stating that

> *"[. . .] ontologies are part of the W3C standards stack for the Semantic Web, in which they are used to specify standard conceptual vocabularies in which to exchange data among systems, provide services for answering queries, publish reusable knowledge bases, and offer services to facilitate interoperability across multiple, heterogeneous systems and databases."*

2.4.2.1 will explain the first formal language describing ontologies which handles the data exchange portion of above citation – the Resource Description Framework (RDF).

2.4.2.1 Resource Description Framework – RDF

The Resource Description Framework (RDF) is a standard by the W3C using XML (see subsection 2.4.1) syntax *"[. . .] for data interchange on the Web"* [28]. Furthermore the W3C states that

> *"RDF extends the linking structure of the Web to use URIs to name the relationship between things as well as the two ends of the link (this is usually referred to as a "triple"). Using this simple model, it allows structured and semi-structured data to be mixed, exposed, and shared across different applications.*
> *This linking structure forms a directed, labeled graph, where the edges represent the named link between two resources, represented by the graph nodes. This graph view is the easiest possible mental model for RDF and is often used in easy-to-understand visual explanations."*

To represent such graphs RDF is using RDF triples – explained in the *RDF/XML Syntax Specification*[29] describing the structure of RDF – which is composed of (as illustrated in figure 2.12) a subject node, predicate and an object node which means that an object describes a subject because they are related in some way (predicate). All three components can be RDF URI references but only the object can be a literal, too.

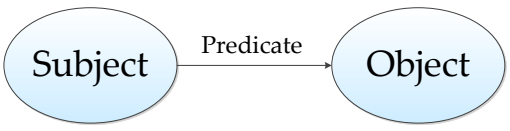


Figure 2.12: *RDF Structure (modified after (Klyne & Carroll, 2004))*

[28] http://www.w3.org/RDF/
[29] (Klyne & Carroll, 2004)

In summary it can be concluded that RDF

- is made for data interchange on the web,
- is using URIs,
- handles relationships between things,
- and forms a directed, labeled graph (i.e. its linking structure).

In its documentation of the RDF Primer (Manola & Miller, 2004) the W3C is giving an example for the usage of RDF and the representation of RDF as a graph. It specifies a resource with these statements which are altered to fit in the marine domain:

(1) There is a Person identified by http://www.baw.de/kontakt/RL
(2) whose name is Rainer Lehfeldt
(3) whose email address is rainer.lehfeldt@baw.de
(4) and whose title is Dr.

```
1  <?xml version="1.0"?>
2  <rdf:RDF
3    xmlns:rdf="http://www.baw.de/rdf/syntax#"
4    xmlns:contact="http://www.baw.de/kontakt#">
5    <contact:Person rdf:about="http://www.baw.de/kontakt/RL">
6      <contact:fullName>Rainer Lehfeldt</contact:fullName>
7      <contact:mailbox rdf:resource="rainer.lehfeldt@baw.de"/>
8      <contact:personalTitle>Dr.</contact:personalTitle>
9    </contact:Person>
10 </rdf:RDF>
```

Listing 2.1: *RDF/XML describing Rainer Lehfeldt (inspired by (Manola & Miller, 2004))*

(1) is the subject and it is identified by an URI while the objects describing this subject are (2) to (4). The subjects are also containing the predicates: *whose name is*, *whose email address is* and *whose title is*. Figure 2.13 shows that the predicates (arrows in the figure) also have URIs. Besides the three subjects (yellow) and the object (blue) there is also a type in the figure (red). With the predicate shown in the figure it specifies that the subject is of type http://www.baw.de/kontakt#Person. The RDF/XML representation corresponding to figure 2.13 is shown in listing 2.1.

For the representation in RDF triples subjects are defined as underlined, predicates as *italics* and objects as **bold**. The syntax of this representation[30] is: subject *predicate* **object** and when applied to the example describing Rainer Lehfeldt the result looks like this:

(1) http://www.baw.de/kontakt/RL *http://www.baw.de/kontakt#* **contact**
(2) contact#me *http://www.baw.de/rdf/syntax#type* **contact#Person**
(3) contact#me *contact#fullName* **'Rainer Lehfeldt'**
(4) contact#me *contact#personalTitle* **'Dr.'**
(5) contact#me *contact#mailbox* **rainer.lehfeldt@baw.de**

(1) is a declaration for the URI http://www.baw.de/kontakt/RL which is used in (almost) every subject and predicate so that **contact** can be used instead of this rather long URI. (2) defines that the subject contact#me is of the type **contact#Person** and (3) to (5) define the object's title, name and mail address of the subject by the predicates *contact#personalTitle*, *contact#fullName* and *contact#mailbox*.

[30] similar to the N-Triples notation

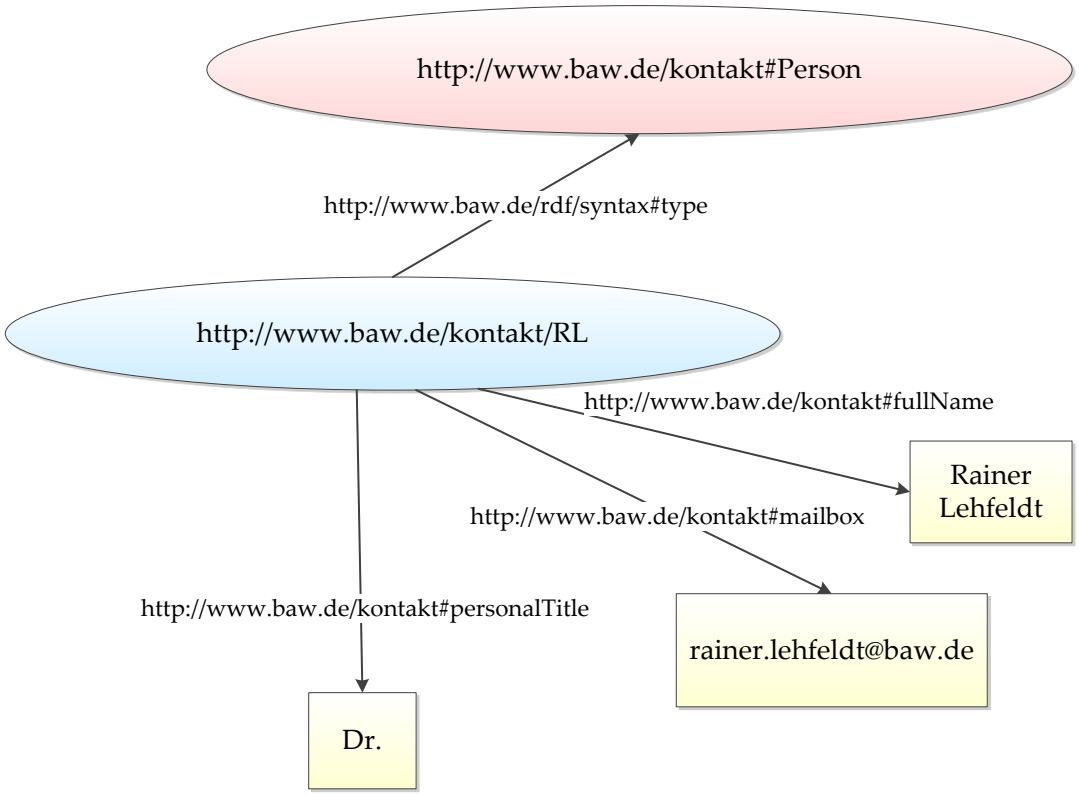


Figure 2.13: *An RDF Graph describing Rainer Lehfeldt (inspired by (Manola & Miller, 2004))*

2.4.2.2 Simple Knowledge Organisation System – SKOS

Another standard for the representation of controlled vocabularies that is built upon the Resource Description Framework is the Simple Knowledge Organisation System (SKOS, also specified by the W3C) which[31]:

> *"[. . .] is a formal language for representing controlled structured vocabularies such as thesauri or classification schemes."*

Because SKOS is an application of RDF it[32] *"[. . .] can be used to express the content and structure of a concept scheme as an RDF graph."* A very simple graph is shown in figure 2.14 which is the first step in the example now to be built[33] which will model certain aspects about the term *beach*. The figure shows the definition of a resource (i.e. a term) called `ex:beach` which is of `rdf:type` `skos:Concept`. Listing 2.2 shows the RDF/XML syntax representing the figures graph. The listing was just used to show the `rdf:type` usage. The remaining examples will use the shortened form shown in listing 2.3.

```
1 <rdf:RDF
2   xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#">
3
4   <rdf:Description rdf:about="http://example.net/concepts#beach">
5     <rdf:type rdf:resource="http://www.w3.org/2004/02/skos/core#Concept"/>
6   </rdf:Description>
7 </rdf:RDF>
```

Listing 2.2: *RDF/XML syntax of the SKOS concept beach*

[31] (Miles, 2006)
[32] (Miles et al., 2005)
[33] cf. (Miles & Brickley, 2005)

```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:skos="http://www.w3.org/2004/02/skos/core#">

  <skos:Concept rdf:about="http://example.net/concepts#beach"/>
</rdf:RDF>
```

Listing 2.3: *RDF/XML syntax of the SKOS concept beach (shortened)*

Throughout the example the prefix `skos:` is used to abbreviate the URI http://www.w3.org/2004/02/skos/core# meaning that e.g. `skos:prefLabel` written out is http://www.w3.org/2004/02/skos/core#prefLabel. The two further prefixes used are `rdf:` which abbreviates the URI http://www.w3.org/1999/02/22-rdf-syntax-ns# and `ex:` which simulates an own URI and is defined as http://example.net/concepts.

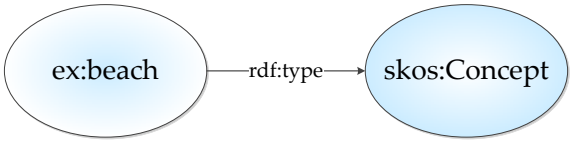


Figure 2.14: *An RDF Graph defining the SKOS concept beach (roughly based on (Miles & Brickley, 2005))*

SKOS classes

The first element (`skos:Concept`) of the SKOS data model was introduced in the short example. A `skos:Concept` is a SKOS class[34] which[35] *"[. . .] can be viewed as an idea or notion; a unit of thought."* According to the SKOS reference (Miles & Bechhofer, 2009) there are three further SKOS classes aside from `skos:Concept` (but to keep things as short as possible only `skos:Concepts` will be explained in more detail):

- `skos:Collection` – labeled and/or ordered groups of SKOS concepts
- `skos:OrderedCollection` – ordered group with meaningful ordering
- `skos:ConceptScheme` – aggregation of one or more SKOS concepts (used for data from unknown or external sources)

SKOS properties

There is a range of properties which can be assigned to a `skos:Concept`:

(1) Labelling properties
(2) Documentation properties
(3) Semantic relationships

(1) Labelling[36] *"[. . .] means assigning some sort of token to a resource, where the token is intended to be used to denote (label) the resource in natural language discourse and/or in representations intended for human consumption."* SKOS offers five properties:

(1) `skos:prefLabel`
(2) `skos:altLabel`
(3) `skos:hiddenLabel`
(4) `skos:prefSymbol`
(5) `skos:altSymbol`

The symbolic labelling (4) and (5) labels a concept with an image. More important for the usage in this thesis are the labeling properties (1) to (3). (1) and (2)[37] *"[. . .] allow you to assign preferred*

[34] Note that (Lord, 2010) used *class* synonymous with *concept*.
[35] (Miles & Bechhofer, 2009)
[36] (Miles & Brickley, 2005)
[37] (Miles & Brickley, 2005)

and alternative lexical labels to a resource." A language tag can be applied to these types of labels as figure 2.15 and listing 2.4 show. Due to this option a multilingual thesaurus can be built with SKOS and later in a web application users see the labels for terms according to the language they configured to use. Listing 2.4 also displays `skos:hiddenLabels` (3) which can be accessed by applications (e.g. for text-based indexing and search functions) but are not visible otherwise. These can be used for typos for example, so that users find certain terms even if they mistype them.

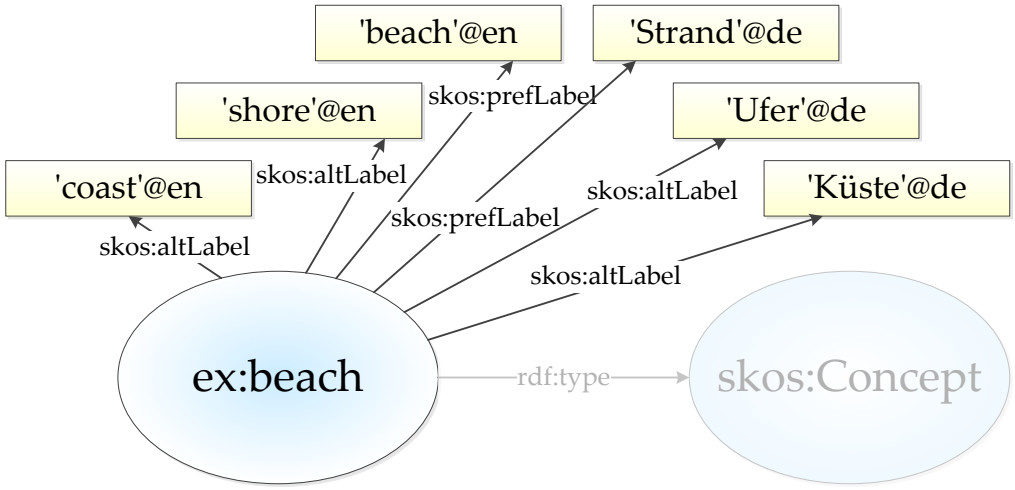


Figure 2.15: *An RDF Graph labelling the SKOS concept beach (roughly based on (Miles & Brickley, 2005))*

```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:skos="http://www.w3.org/2004/02/skos/core#">
  <skos:Concept rdf:about="http://example.net/concepts#beach">
    <skos:prefLabel xml:lang="en"> beach </skos:prefLabel>
    <skos:altLabel xml:lang="en"> shore </skos:altLabel>
    <skos:altLabel xml:lang="en"> coast </skos:altLabel>
    <skos:prefLabel xml:lang="de"> Strand </skos:prefLabel>
    <skos:altLabel xml:lang="de"> Ufer </skos:altLabel>
    <skos:altLabel xml:lang="de"> Küste </skos:altLabel>
    <skos:hiddenLabel> beachh </skos:hiddenLabel>
    <skos:hiddenLabel> caost </skos:hiddenLabel>
    <skos:hiddenLabel> shor </skos:hiddenLabel>
  </skos:Concept>
</rdf:RDF>
```

Listing 2.4: *RDF/XML syntax of the SKOS concept beach with multilingual labels and hidden typos*

The labels discussed so far are *lexical entities* which means that they are more or less string literals. Because of that the labels are not objects themselves and thus cannot be described further with metadata. To add information about the labels – such as who was the author of a particular label or when was the label last updated – the W3C built an extension for SKOS – SKOS eXtension for Labels (**SKOS-XL**) – which

> *"[. . .] defines an extension for the Simple Knowledge Organization System, providing additional support for describing and linking lexical entities."*[38]

The "new" labels are `skosxl:prefLabel`, `skosxl:altLabel` and `skosxl:hiddenLabel` which are instances of the class `skosxl:Label`. Instances of this class have a `skosxl:literalForm` which holds the label of the concept and on top of that any additional properties a user wants/needs. The usage of the additional properties `:lastEdited` and `:myCustomProperty` of the concept *beach* is shown in listing 2.5.

[38] (Miles & Brickley, 2009)

```
@prefix skos:    <http://www.w3.org/2004/02/skos/core#> .
@prefix skosxl: <http://www.w3.org/2008/05/skos-xl#> .
@prefix:         <http://www.example.com/demo#> .
@prefix rdf:     <http://www.w3.org/2000/01/rdf-schema#> .
@prefix xsd:     <http://www.w3.org/2001/XMLSchema#> .

:concept234 rdf:type skos:Concept ;
  skosxl:prefLabel :label1 .

:label1 rdf:type skosxl:Label ;
  :lastEdited "2011-02-05T10:21:00"^^xsd:dateTime ;
  :myCustomProperty 2.71828 ;
  skosxl:literalForm "beach"@en-GB .
```

Listing 2.5: *Turtle syntax of the SKOS concept beach with added SKOS-XL properties (modified after (DuCharme, 2011))*

(2) Documentation In addition to labels terms can be described more in-depth using documentation properties. SKOS is offering the following seven to add human-readable content to concepts[39]:

- `skos:note` – general documentation for any purpose
  - `skos:definition` – complete explanation of the meaning
  - `skos:scopeNote` – information about what is or is not included within the meaning (scope)
  - `skos:example` – show exemplary use
  - `skos:historyNote` – e.g. reflect or describe changes of a term
  - `skos:editorialNote` – e.g. reminders of future editorial work
  - `skos:changeNote` – document changes

Figure 2.16 and listing 2.6 are showing the usage of two of these documentation properties. The concept *beach* has a definition in English and in German. Because the English definition may not be good, for instance, there is also a `skos:editorialNote` that reminds the user(s) to come up with a better definition.

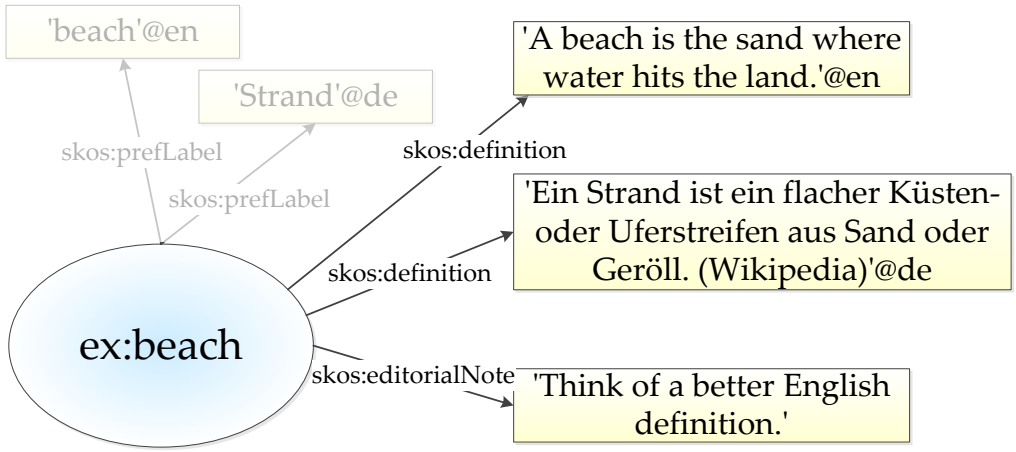


Figure 2.16: *An RDF Graph documenting the SKOS concept beach (roughly based on (Miles & Brickley, 2005))*

[39] cf. (Miles & Brickley, 2005)

```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:skos="http://www.w3.org/2004/02/skos/core#">
  <skos:Concept rdf:about="http://example.net/concepts#beach">
    <!-- labels are left out -->
    <skos:definition xml:lang="en"> A beach is the sand where water hits the land. </
        skos:definition>
    <skos:editorialNote> Think of a better English definition. </skos:editorialNote>
    <skos:definition xml:lang="de"> Ein Strand ist ein flacher Küsten- oder
        Uferstreifen aus Sand oder Geröll. (Wikipedia) </skos:definition>
  </skos:Concept>
</rdf:RDF>
```

Listing 2.6: *RDF/XML syntax documenting the SKOS concept beach*

(3) Semantic Relationships[40] are used to show that a concept relates to another concept by providing a link between them. These relations can be either hierarchical or associative. Figure 2.17 and listing 2.7 are showing the existence of a hierarchical relationship between the concept *beach* and *beach scarp* where *beach* is the superordinate (`skos:broader`) term in relation to *beach scarp* which in turn is a subordinate (`skos:narrower`) term compared to *beach*. The figure and listing are also showing two `skos:related` relationships which are associative and the last type of semantic relationships. In the example there are relations from beach to sand and water and vice versa which is pretty easy to imagine.

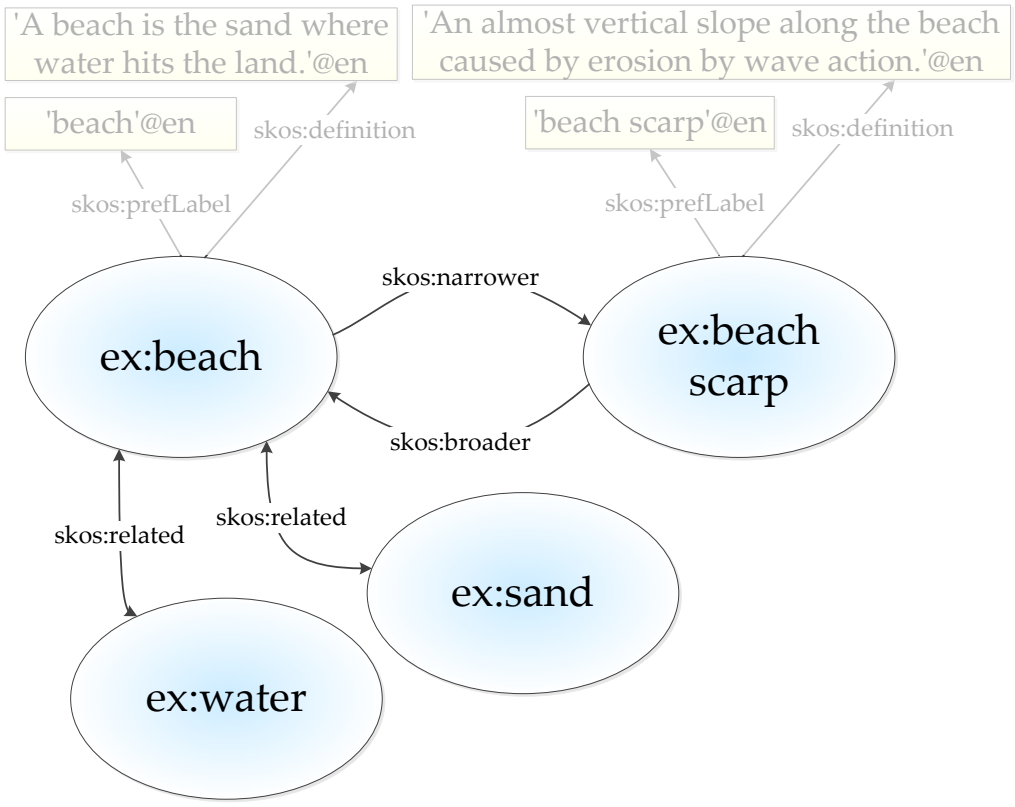


Figure 2.17: *An RDF Graph showing the semantic relationships in relation to beach (roughly based on (Miles & Brickley, 2005))*

[40] cf. (Miles et al., 2005, p. 4) and (Miles & Bechhofer, 2009)

```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:skos="http://www.w3.org/2004/02/skos/core#">
  <!-- labels etc. are left out -->
  <skos:Concept rdf:about="http://example.net/concepts#beach">
    <skos:narrower rdf:resource=www.example.com/concepts#beach_scarp"/>
    <skos:related rdf:resource="www.example.com/concepts#water"/>
    <skos:related rdf:resource="www.example.com/concepts#sand"/>
  </skos:Concept>
  <skos:Concept rdf:about="example.net/concepts#beach_scarp">
    <skos:broader rdf:resource="www.example.com/concepts#beach"/>
  </skos:Concept>
</rdf:RDF>
```

Listing 2.7: *RDF/XML syntax showing the semantic relationships in relation to beach*

2.5 Directives in the marine and SDI domain

Important drivers for the development of a marine SDI are directives. In general there is a close connection between SDIs and directives: SDIs support administrative activities and directives are part of or influence said administrative activities. For European countries like Germany these directives are on the European level and are thus legislated by the European Union. The main directives affecting the marine domain are the Water Framework Directive (WFD, subsection 2.5.2) and the Marine Strategy Framework Directive (MSFD, subsection 2.5.3) but some of the annexes of the INSPIRE directive are also important for the marine domain (as subsection 2.1.5 on page 20 already outlined). However, they also have counterparts in German law or existing German laws were adjusted to meet their requirements for Germany and its federal states. Examples include Meeresstrategie-Rahmenrichtlinie (MSRL respectively MSFD), Wasserrahmenrichtlinie (WRRL respectively WFD), Fauna-Flora-Habitat-Richtlinie (FFH-RL respectively Habitats Directive) and Vogelschutzrichtlinie (VS-RL respectively Birds Directive). Other legislation and directives include the national law on access to spatial data (GeoZG), the Environmental Information Act (UIG) as well as the Water Information System for Europe (WISE), Baltic Marine Environment Protection Commission (HELCOM), Convention for the Protection of the Marine Environment of the North-East Atlantic (OSPAR), EU Shared Environmental Information System (SEIS) and Agenda 21 (United Nations).

2.5.1 INSPIRE

As subsection 2.1.4 pointed out, INSPIRE is a European directive on which a regional SDI (for Europe) will be built. This SDI is located at the regional level because it covers all the EU member states and relies on their national SDIs. Furthermore, INSPIRE is an example for a legally enforced SDI because it is a legal act (directive 2007/2/EC) of the Council of the European Union and the European Parliament. INSPIRE focuses on environmental policy and aims at strengthening the availability and accessibility of data overcoming barriers such as incompatibility and inconsistency (already outlined in section 2.1).

The challenge of achieving these goals is that INSPIRE is built by 27 different countries (and therefore has to support more than 23 languages) with very different information systems, professional and cultural practices. The scope of INSPIRE is defined by 34 themes which fall into three categories or annexes where the first two focus on "fundamental datasets" respectively reference data such as *coordinate reference systems*, *geographical names* and *elevation* while annex III covers data for environmental analysis and impact assessment such as *environmental monitoring facilities* and *sea regions* (Craglia, 2010b).

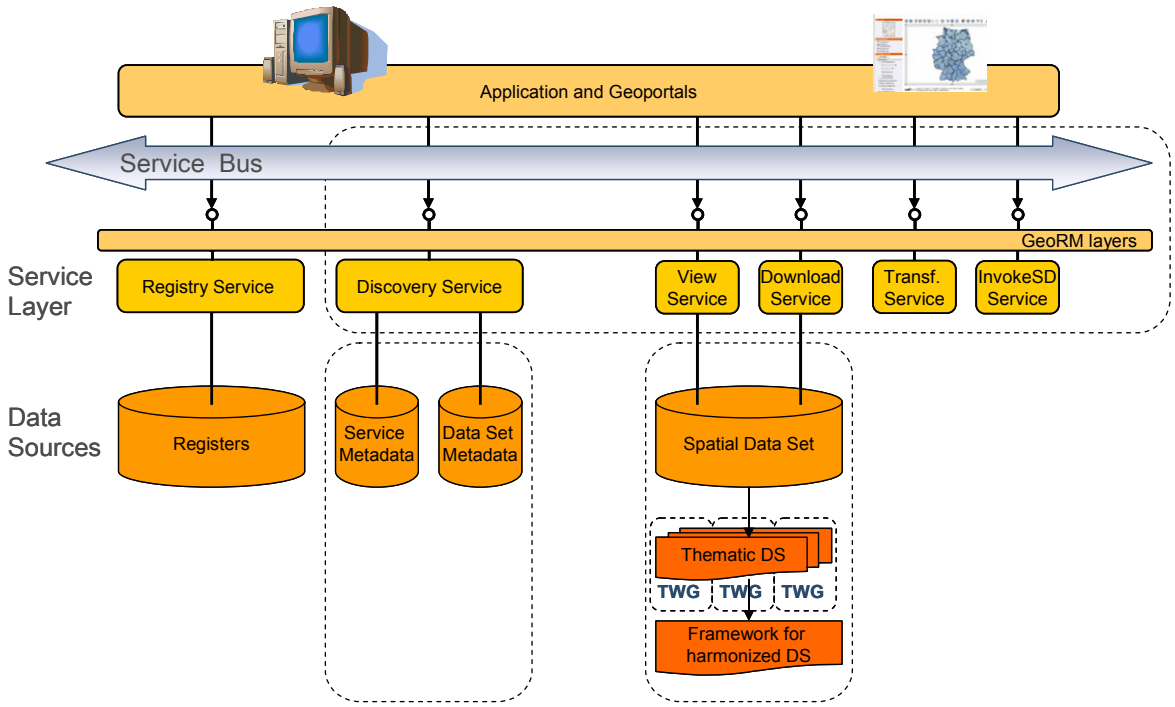


Figure 2.18: *INSPIRE technical architecture overview (INSPIRE Network Services Drafting Team, 2007, p. 4)*

INSPIRE architecture and services The architecture of INSPIRE is depicted in figure 2.18. This figure also shows the core component respectively resource: spatial data in spatial data sets. All other components (metadata, services and so on) are only needed to find, use, interpret or access said spatial data. Data access is possible through services that are described by service metadata (service descriptions) (INSPIRE Network Services Drafting Team, 2007, p. 4). There are six service types used to offer access to data and metadata:

- Registry Services,
- Discovery Services,
- View Services,
- Download Services,
- Transformation Services and
- Invoke Spatial Data Services.

Registry services allow access to registers that describe data so that it is processable and interpretable. However, there are no Implementing Rules for registry services. In contrast to that *discovery services* seem to be of higher importance since Implementing Rules exist for this service type. A discovery service offers *"[. . .] search for spat ial data sets and services on the basis of the content of the corresponding metadata and to display the cont ent of the metadata."*[41] After a system or a user found spatial data sets via a discovery service the data sets can be viewed or downloaded through *view services* and *download services*. *Transformation services* are Coordinate Transformation Services ultimately and can either run on view or download services or as their own service instance. *Invoke spatial data services* allow *"[. . .] defining both the data inputs and data outputs expected by the spatial service and define a workflow or service chain combining multiple services. It also allows the definition of a web service interface managing and accessing (executing) workflows or service chains."*[42]

[41] (INSPIRE Network Services Drafting Team, 2008, p. 9)
[42] (INSPIRE Network Services Drafting Team, 2008, p. 11)

INSPIRE and marine data According to (Korduan, 2013) more than half of the 34 INSPIRE themes affect marine data in a broader or narrower sense. The themes which most probably involve marine data are:

- Geographical names (I)
- Administrative units (I)
- Transport networks (I)
- Hydrography (I)
- Protected sites (I)
- Agricultural and aquaculture facilities (III)
- Area management/restriction/regulation zones and reporting units (III)
- Bio-geographical regions (III)
- Environmental monitoring facilities (III)
- Energy resources (III)
- Habitats and biotopes (III)
- Land use (III)
- Mineral resources (III)
- Natural risk zones (III)
- Oceanographic geographical features (III)
- Sea regions (III)
- Species distribution (III)
- Statistical units (III)
- Utility and Government services (III)

Using *Hydrography* as an example the marine context is very obvious as it directly involves marine data which is even clearer when looking at the structure of its application schema which shows that it is subdivided into[43]:

- **Network**
  - *Watercourse*
  - *Hydro Node*
- **Physical Waters**
  - *Wetland*
  - *Shore*
  - *Drainage Basin*
- **Reporting**
  - *WFDCoastalWater*
  - *WFDSurfaceWaterBody*

The list shows that the entire feature types of the application schema *Reporting* have the prefix "WFD" indicating Water Framework Directive. The inclusion of these WFD reporting objects make it possible to link them to related hydrographic objects in other application schemas meaning a *WFDRiver* may be related to a *Watercourse* (Lekkerkerk et al., 2010, pp. 53).

2.5.2 Water Framework Directive (WFD)

The Water Framework Directive (WFD, directive 2000/60/EG) aims at achieving a "good status" for rivers, lakes, groundwater bodies and coastal waters in Europe with an implementation deadline of 15 years which means that it aims at achieving said "good status" by 2015. The WFD provides a common framework for European water policy by linking physical planning with water resource planning and defining that water quality goes hand in hand with emission controls and groundwater protection. Apart from focussing on water quality the WFD centres on improving biodiversity (Flasbarth, 2011) and (Kaika, 2003, pp. 299).

INSPIRE was able to include WFD reporting objects because its implementation began on December, 22nd 2000. The WFD is described by (Kaika, 2003) as a

> *"[. . .] legally binding policy that provides a common framework for water management and protection in Europe and that promises to transform the European water sector."*

[43] exemplary feature types given in *italics*

2.5.3 Marine Strategy Framework Directive (MSFD)

Article 1.1 of directive 2008/56/EC[44] (Marine Strategy Framework Directive) summarizes the directive very well:

> *"This Directive establishes a framework within which Member States shall take the necessary measures to achieve or maintain good environmental status in the marine environment by the year 2020 at the latest."*

To achieve the aim of a marine GES (good environmental status) the directives proposes strategies which *"[. . .] protect and preserve the marine environment, prevent its deterioration or, where practicable, restore marine ecosystems in areas where they have been adversely affected"* and *"[. . .] prevent and reduce inputs in the marine environment [. . .] to ensure that there are no significant impacts on or risks to marine biodiversity, marine ecosystems, human health or legitimate uses of the sea."*

The initial assessment in conjunction with the definition what a GES is and the foundation of environmental targets and associated indicators was completed on July, 15th 2012. By July, 15th 2014 a monitoring programme for on-going assessment and regular updating of said targets has to be implemented. The programme of measures which aims at achieving or maintaining a GES has to be developed by 2015 and has to be in operation by 2016 (Zampoukas et al., 2012, p. 26).

The assessment of the reports which have to be prepared for the two programmes (monitoring and measures) are based on eleven descriptors – which are further subdivided into criteria and indicators – which are defined in Annex I[45]:

(1) Biodiversity
(2) Invasive species
(3) Commercially exploited species
(4) Food webs
(5) Eutrophication
(6) Sea-floor integrity
(7) Hydrographical conditions
(8) Contaminants and pollution effects
(9) Contaminants in fish and other seafood
(10) Marine litter
(11) Underwater noise/energy

Indicators for descriptor 5 (Eutrophication) are *Nutrients concentration in the water column* and *Chlorophyll concentration in the water column* for instance.

2.6 Conclusions

Section 2.1 showed that interoperability is a prerequisite for SDI development because an SDI is all about systems and users working together, i.e. exchanging and working on data. Interoperability means that systems are able to communicate with each other or simply that users are able to use data that was given to them. Interoperability is achieved through standards and specifications. Apart from standards like ISO 19115 and (open) specifications like OGC's WMS these can also be de facto standards like Esri's Shapefile that enables users to exchange and use data. Successful SDI development has to be based on standards (ISO) and specifications (OGC) because they make interoperability possible.

A base for SDI design may be the use of a reference model which can aid SDI development by constituting decisions and offering a guideline. Another advantage of a reference model is its division into several aspects respectively parts. Thus it follows the *divide and conquer* approach and makes certain things easier to handle, discuss and implement.

Particular aspects in SDI development are affected (or even promoted) by directives which are also reflected in a reference model. Directives (may) propose special requirements regarding data and

[44](MSFD, 2008, p. 24)
[45]shortened version taken from (Zampoukas et al., 2012, pp. 13)

metadata harmonization, metadata annotation and so on that have to be taken into account when developing a SDI. These requirements have to be known from the beginning and thus are constituted in the reference model.

3 Existing approaches and established systems

> *"Incorporation of marine and coastal regions within global, national and regional spatial data infrastructures (SDIs) should bring substantial additional benefits of integration, standardisation and interoperability of technologies, enabling better policy formulation, monitoring and enforcement [. . .].*
> *Creating marine and coastal zone SDIs, however, is a far from trivial challenge."*
>
> (Bartlett et al., 2004, p. 2)

As already outlined Germany is currently developing a MSDI which will be introduced firstly (see section 3.1). However, spatial data infrastructures in general and specifically marine SDIs are not new and thus there are already existing systems. That is why section 3.2 introduces several existing marine SDIs worldwide. No additional systems are introduced because these suffice as base for the following research (see inter alia chapter 4).

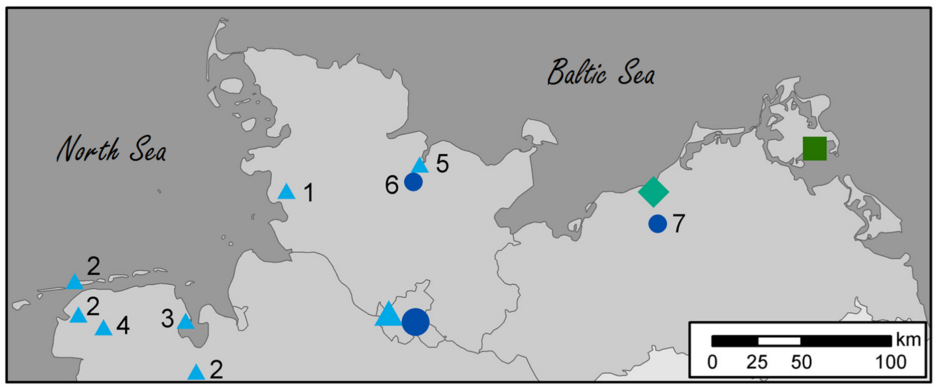
Section 3.3 looks at the usage of reference models in order to construct SDIs. Furthermore section 3.4 describes the existing marine vocabularies in Germany which are an important component of a SDI. To convert these lists and manage vocabularies respectively thesauri online section 3.5 looks at conversion tools and online thesaurus management tools.

3.1 Germany: MDI-DE

The aim of the marine data infrastructure Germany (MDI-DE) is to integrate existing technical developments (NOKIS – a distributed metadata management system in Germany – and the spatial data infrastructure of the German Federal Maritime and Hydrographic Agency [GDI-BSH]) thus merging information concerning the fields coastal engineering, hydrography and surveying, protection of the marine environment, maritime conservation, regional planning and coastal research. The project was funded by the Federal Ministry of Education and Research (BMBF) from 2010 till 2013. The funded parties and their sub projects (SPs) in this project were:

- Federal Waterways Engineering and Research Institute (Bundesanstalt für Wasserbau BAW, SP1 – "coastal engineering and coastal water protection"),
- German Federal Maritime and Hydrographic Agency (Bundesamt für Seeschifffahrt und Hydrographie BSH, SP2 – "protection of the marine environment"),
- German Federal Agency for Nature Conservation (Bundesamt für Naturschutz BfN, SP3 – "maritime conservation") and
- Professorship for Geodesy and Geoinformatics at Rostock University (GG, SP4 – "scientific accompanying research").

Figure 3.1 shows the projects participants from federal and state institutions with responsibilities in the North and Baltic Sea and their location in detail.



Legend

SP1: Coastal engineering and coastal water protection

Principal applicant
Federal Waterways Engineering and Research Institute

Project participants
1 Authority for coastal protection, national parks and marine protection in Schleswig-Holstein
2 Authority for water management, coast protection and nature conservation in Lower Saxony
3 National Park Office of the Lower Saxony Wadden Sea
4 Federal Administration of Waterways and Navigation – Directorate Northwest
5 Federal Administration of Waterways and Navigation – Directorate North

SP2: Protection of the marine environment

Principal applicant
Federal Maritime and Hydrographic Agency

Project participants
6 State office for agriculture, environment and rural areas
7 State office for environment, conservation and geology

SP3: Maritime conservation

Federal Agency for Nature Conservation

SP4: Scientific accompanying research

Professorship for Geodesy and Geoinformatics, University of Rostock

Figure 3.1: *Project participants, sub projects (SP) and their locations (translated from and modified after (Lehfeldt & Melles, 2011, p. 107))*

As can be seen in figure 3.2 this undertaking is related to a number of regulations and developments on many administrative levels from which specifications and courses of action derive. On the European level it is the INSPIRE (Infrastructure for Spatial Information in the European Community) initiative as well as the Marine Strategy Framework Directive (MSFD), the Water Framework Directive (WFD) and Natura2000 with their regulation counterparts for Germany and its federal states (MSRL, WRRL, FFH-RL, VS-RL).

The main components of MDI-DE are shown in the overview that figure 3.3 gives. The outer ring shows (some of) the infrastructure nodes that were respectively will be built by project participants (e.g. LUNG) and partners (e.g. AWI). The red inner ring highlights the main componts MDI-DE is made of: portal, services, thesaurus etc. In its innermost circle the figure also accentuates the merging of NOKIS and the SDI of the German Federal Maritime and Hydrographic Agency (GDI-BSH).

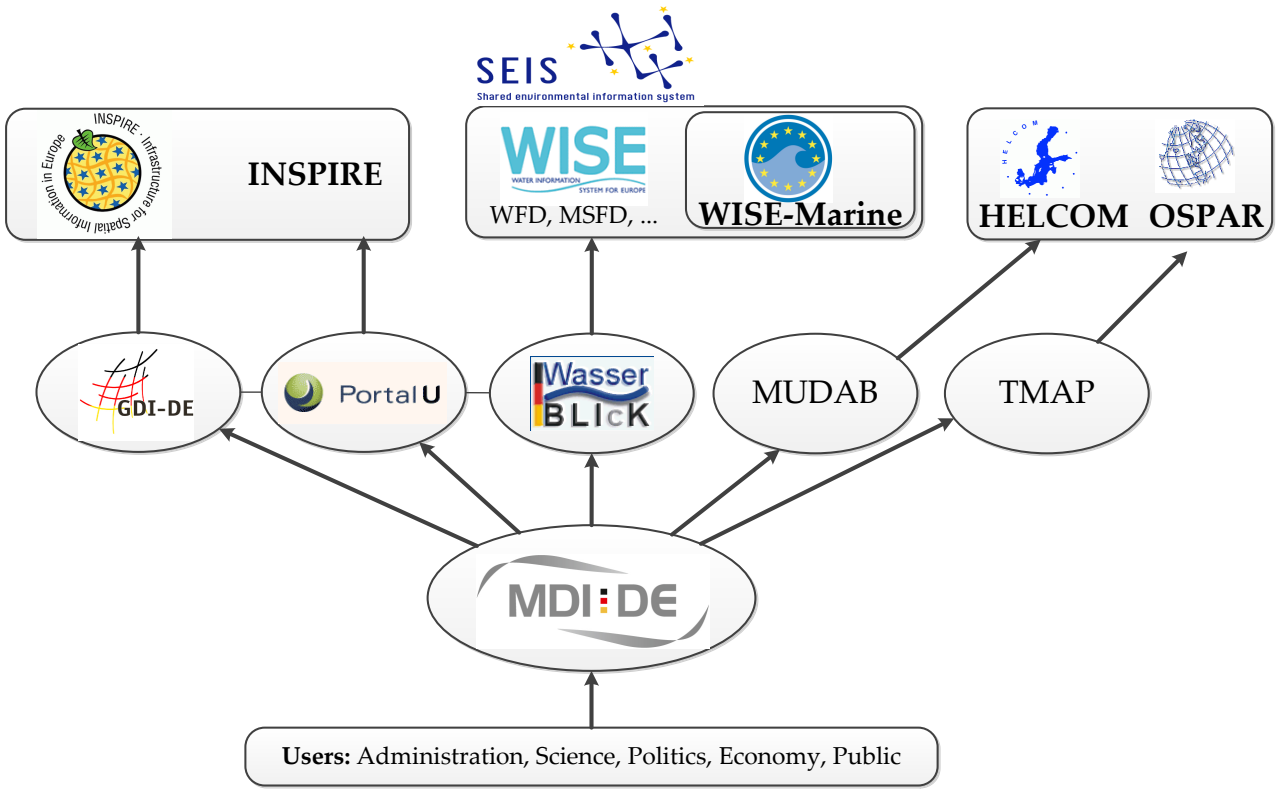


Figure 3.2: *MDI-DE in the scope of actions from German and European initiatives (translated from and modified after (Kohlus & Reimers, 2010, p. 122))*

All the aspects mentioned so far (relations to directives, components etc.) are represented by the nine working groups (WGs) of MDI-DE:

(1) reference model
(2) evaluation of existing data sets
(3) infrastructure nodes
(4) portal
(5) MSFD activities
(6) data harmonization and interoperability
(7) metadata modelling
(8) Sensor Observation Service (SOS)
(9) editorial activities

The project's outcomes are available under www.mdi-de.org. Documentation in form of reports and information on the final publication in *"Die Küste"* can also be found there. In the future MDI-DE will be continued and maintained by federal and state institutions based on the *agreement on cooperation for design and development of software for environmental information systems* (VKoopUIS).

3.2 International MSDIs

In order to learn from other initiatives several marine SDI approaches worldwide will be introduced. Subsection 3.2.1 describes two approaches from Australia which are part of the Australian Spatial Data Infrastructure (ASDI): the Australian Marine Spatial Information System (AMSIS) and Integrated Marine Observing System (IMOS) Ocean Portal. Subsection 3.2.2 introduces three components of the Canadian Geospatial Data Infrastructure (CGDI) also known as "GeoConnections": the Marine Geospatial Data Infrastructure (MGDI), Department of Fisheries and Oceans (DFO) GeoPortal and COINAtlantic (Coastal and Ocean Information Network Atlantic). While these two countries have more than one entry point for marine data Ireland has a single source for marine data – the Marine Irish Digital Atlas (MIDA, subsection 3.2.3). As another example for an European country's MSDI

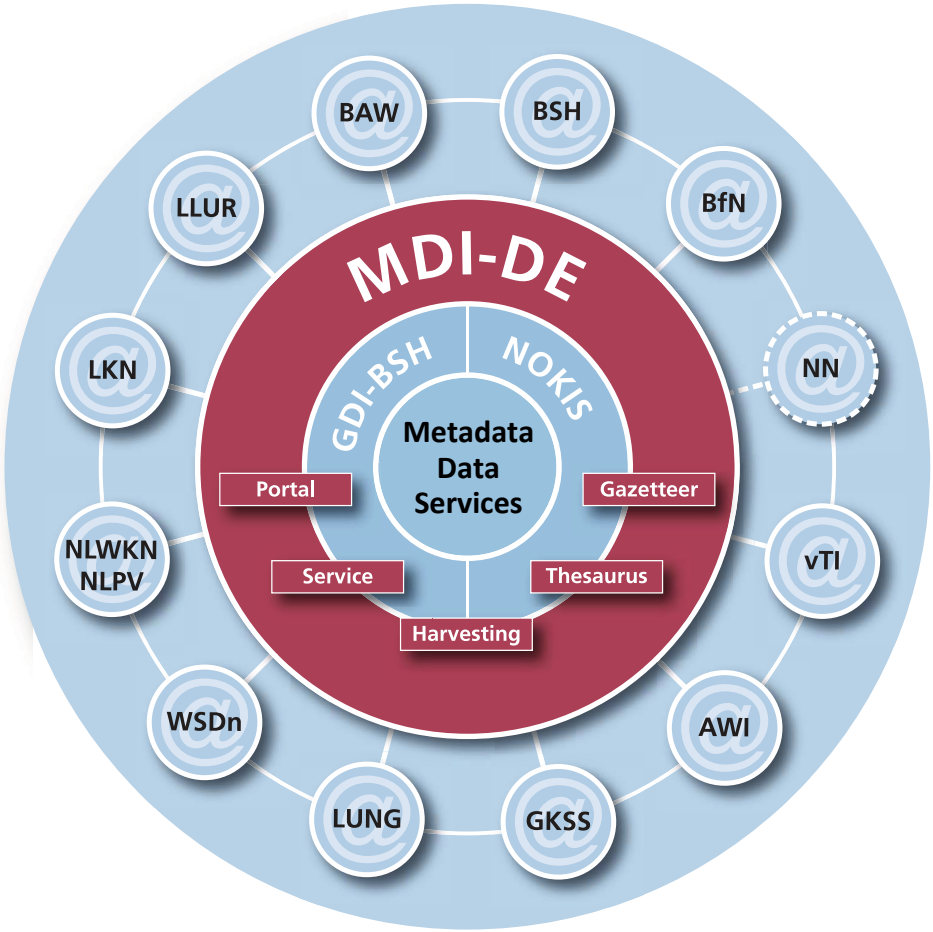


Figure 3.3: *Components of MDI-DE (translated from and modified after (Lehfeldt & Melles, 2011, p. 110))*

the efforts of the United Kingdom will be presented in subsection 3.2.4 – MAGIC/CAMRA and MEDIN. Since the United States of America were and are an impetus in the field of SDI and MSDI development their Coastal and Marine Spatial Planning (CMSP) Registry, Marine Cadastre, Data.gov and Geoplatform will also be introduced in subsection 3.2.5. Later, chapter 4 will evaluate and analyse these approaches to see what can be learned from them or what to avoid.

3.2.1 Australia: ASDI, AMSIS and Ocean Portal

In Australia we mainly find two approaches with the aim to tie in with the Australian Spatial Data Infrastructure (ASDI). One – the Australian Marine Spatial Information System (AMSIS) – focuses primarily on "framework" data (boundaries, cadastre, infrastructure etc.) while the other – the Integrated Marine Observing System (IMOS) Ocean Portal – offers a variety of data mostly from scientific research e.g. biological and climate data.

AMSIS was developed and is maintained by Geoscience Australia (government agency) and – as a web based interactive mapping and decision support tool – offers access to over 80 layers of information in the Australian marine jurisdiction including maritime boundaries, bathymetry, physical and environmental information, legal interests, fisheries and shipping (Dwyer & Wright, 2008).

The Integrated Marine Observing System (IMOS) is a distributed set of equipment and data-information services which, among many other applications, aims at meeting the needs of the research community in Australia. The strategic focus of IMOS is on the 4-dimensional ocean

variability and the impact of major boundary currents on the continental shelf, ecosystems and biodiversity. IMOS is organized as a matrix of nodes and facilities where the facilities deliver data streams which are then used by the nodes and other stakeholders. There are facilities inter alia for bluewater, climate observations, coastal currents, water properties, coastal ecosystems and a biophysical sensor network on the Great Barrier Reef. The data the facilities are producing are made available through the electronic Marine Information Infrastructure (eMII). eMII is based at the University of Tasmania and handles and organises the storage, accessibility, discoverability and means of visualisation of data. All data is freely available from the IMOS Ocean Portal which allows the discovery and usage of the data from all of the facilities (Moltmann et al., 2010) and (Proctor et al., 2010).

3.2.2 Canada: Marine Geospatial Data Infrastructure (MGDI), COINAtlantic and GeoPortal

The national spatial data infrastructure (NSDI) of Canada is called Canadian Geospatial Data Infrastructure (CGDI) but is also known as "GeoConnections" which is the more market-oriented title and is divided into twelve committees respectively nodes. The CGDI[1] *"[. . .] recognizes that governments have a responsibility to make geospatial information available, and to 'play their role' in developing a knowledge economy in response to the needs of citizens, industry and communities in support of the economic, social and environmental well-being."* The CGDI aims at helping users access and integrate said geospatial information by facilitating the infrastructure. Thus the CGDI does not house the spatial data but provides the framework so that various authorities can provide their data through the use of common standards. The CGDI mainly consists of[2]:

- *"the GeoConnections Discovery Portal (GDP), a national search engine that allows providers to catalogue their data sets and users to determine which data sets exist where;*
- *GeoGratis, a national repository where suppliers may place data for free distribution;*
- *GeoBase, a national suite of framework layers coordinated by the Canadian Council on Geomatics that includes place names, a national digital elevation model, a national layer of satellite imagery, a national road network, national geodetic (survey reference) points, and a national layer of administrative boundaries"*

One component of the CGDI is the Marine Geospatial Data Infrastructure (MGDI) which tries[3] *"[. . .] to satisfy the geographic data needs of water-oriented stakeholders."* The development of the MGDI is led by the Marine Advisory Network node which is one of the CGDIs twelve nodes whereupon the Department of Fisheries and Oceans (DFO) and the Canadian Centre for Marine Communications (CCMC) are the key participants of the node (DFO, 2001). The MGDI assists the economic and social needs of Canada's marine regions and assists the management of Canada's water resources (NRC, 2003). As a key partner of both CGDI and MGDI the Department of Fisheries and Oceans (DFO) is developing the (DFO) GeoPortal which is a key component of the MGDI and provides services that enable DFO employees to index and publish their data and additionally find, view and download other spatial data. The GeoPortal does not intend to be a data warehouse but rather acts as a clearing house for marine spatial data by using an open standards-based architecture [BCMSRM03]. Another initiative inside CGDI is COINAtlantic which[4] *"[. . .] has implemented a coastal and ocean information network for the western North Atlantic."* The initiative is led by the Atlantic Coastal Zone Information Steering Committee (ACZISC) and aims at the provision of open access to spatial data to support integrated coastal and ocean management (ICOM) by adopting all standards of and complying with the architecture of the CGDI (Sherin et al., 2009).

[1](Labonte et al., 1998)
[2](Sherin, 2007)
[3](NRC, 2003, p. 1)
[4](Butler et al., 2011)

3.2.3 Ireland: Marine Irish Digital Atlas

The Marine Irish Digital Atlas (MIDA) originally was a three year project by the Coastal & Marine Resources Centre (CMRC) at University College Cork which started in September 2002 and is still enduring. MIDA[5] *"[. . .] aims to be a single source for marine and coastal geospatial information in Ireland"* from numerous data owners for professional and public use. It provides over 140 data layers (and associated metadata) from more than 35 data sources trying to address the needs of the Irish coastal and marine community including marine scientists, administrators, educational establishments and the general public (Dwyer et al., 2003), (O'Dea et al., 2009) and (O'Dea et al., 2007).

3.2.4 United Kingdom: MAGIC/CAMRA and MEDIN

Multi-Agency Geographic Information for the Countryside (MAGIC) is a web map application that combines data on key environmental schemes and designations and which involves six government organisations (Defra [Department for Environment, Food and Rural Affairs], English Heritage, Environment Agency, Marine Management Organisation and two others)[6]. The Coastal and Marine Resource Atlas (CAMRA) is/was managed by the Maritime and Coastguard Agency (MCA) and is hosted as a sub-topic on the MAGIC website[7]. The atlas features a list of priority datasets including important coastal and marine habitats and species, as well as physical geography and relevant infrastructure. The Atlas is a web map tool offering access to a wide range of information on coastal and marine resources (DEFRA, 2006, p. 1).

The UK Marine Environmental Data and Information Network (MEDIN) is offering a framework for marine data management in the UK through clearing up terms and conditions for data use, coordinating marine survey and research activities, defining data specifications for improved data management and of course better access to data. MEDIN is working through seven interlinked work streams (WS) inter alia concerning a network of marine Data Archive Centres (DACs, WS1), data and metadata specifications (WS2) and a web portal (WS3) for example. Another major issue for MEDIN is longevity which means that it investigates the question on how to achieve a sustainable framework in the long run. MEDIN approaches this question by defining regular meetings, archiving the data in DACs, using an agreed-on metadata standard for all its data, keeping the metadata freely available and many more (Charlesworth et al., 2009).

3.2.5 United States of America: CMSP Registry, Marine Cadastre, Data.gov and Geoplatform

For the management of coastal areas and waters in the United States, both national organizations such as the Marine Cadastre or the scientifically oriented National Ocean and Atmospheric Administration (NOAA) are responsible, as well as organizations from individual states. Through the *National Ocean Policy Implementation Plan* the terrestrial and marine spatial data systems in the U.S. were brought together.

Amongst other portals that are on a state or local level (Oregon Coastal Atlas, North Coast Explorer etc.) and other portals (such as nowCOAST and Digital Coast) there are two marine-oriented portals that are of main importance:

- *Coastal and Marine Spatial Planning Data Registry* (CMSP) by the NOAA *"[. . .] is a collection of Web-accessible NOAA geospatial data deemed essential for local, regional, or national-level CMSP processes."*[8]

[5](Strain et al., 2006, p. 13)
[6]http://magic.defra.gov.uk/About_MAGIC.htm
[7]http://magic.defra.gov.uk/camra.html
[8]http://egisws02.nos.noaa.gov/cmspgisdataregistry/

- *Marine Cadastre* which *"[. . .] is an integrated marine information system that provides ocean data, offshore planning tools, and technical support to the offshore renewable energy community"*[9] developed by NOAA's Coastal Services Center and the Department of the Interior's Bureau of Ocean Energy Management (BOEM)

But marine data has also been incorporated into wider platforms or portals whereat the two most prominent examples are:

- *Geospatial Platform* by the Federal Geographic Data Committees (FGDC) *"[. . .] is an Internet-based capability providing shared and trusted geospatial data, services, and applications for use by the public and by government agencies and partners to meet their mission needs."*[10]
- *Data.Gov* by the Executive Branch of the Federal Government *"[. . .] increases the ability of the public to easily find, download, and use datasets that are generated and held by the Federal Government."*[11]

3.3 Reference models for SDIs

Chapter 5 will define a reference model for the German marine spatial data infrastructure (MDI-DE) in section 5.2. After section 2.3 pointed out a standard which could be used for the definition of a reference model (RM-ODP) this section will show examples of existing reference models for various kinds of SDIs which section 5.2 will use as an additional base for its definition.

(Tóth et al., 2012, p. 28) defines the usage of a reference model for SDIs as:

> *"The reference model states where standards are applicable and how they should be used for developing the data component of the SDI. [. . .] The reference model also lists the types of information technology services that might be used for accessing, processing, and sharing geographic data and related information in the infrastructure."*

After this definition (Tóth et al., 2012) takes ISO standard 19101 (Geographic Information – Reference model) as an example for a reference model. This standard was already mentioned when discussing the Reference Model of Open Distributed Processing (RM-ODP) in section 2.3 on page 29. RM-ODP was used as the base for reference models of SDIs like Australia's Water Resources Observation Network (WRON) which is described in subsection 3.3.2. There is also a variety of reference models which are not based on RM-ODP worth looking at because they include interesting concepts. Apart from two examples from two German federal states – North Rhine-Westfalia in 3.3.1.1 and Saxony in 3.3.1.2 – this section is also concerned with the Digital Earth Reference Model in subsection 3.3.3. Since the SDIs of the two states are located inside the national spatial data infrastructure for Germany (GDI-DE) its reference model will be handled first.

3.3.1 Selected reference models in Germany

The GDI-DE is the national spatial data infrastructure for Germany which aims at improving the provision and usage of distributed spatial data originating from different administration levels, business as well as science. In order to achieve this the GDI-DE tries to simplify information, communication and transaction processes and is breaking down this architecture into three layers (applications, services and data storage) and four sub models. In contrast to RM-ODP the reference model of the GDI-DE is using sub models instead of viewpoints but the sub models are very similar to the viewpoints because they are offering different respectively more detailed views on the whole system (Grohmann & Stahl, 2010).

[9] http://www.marinecadastre.gov/MMC%20Pages/about.aspx
[10] http://www.geoplatform.gov/overview-page
[11] http://www.data.gov/about

The *business model* occupies itself with the components of the GDI-DE and economic efficiency and performance of these components. Economic efficiency and performance are important because GDI-DE will provide its components (e.g. services) to smaller agencies which cannot afford to publish their data for INSPIRE on their own.

The main task of the *architecture model* is to develop the existing architecture further. This starts from business objectives from which tasks for administrations derive. The architecture model revolves around business processes from which requirements and thus specific needs arise. To satisfy these needs the system offers services which in turn need an IT infrastructure. Infrastructure components such as hardware, network, databases and what is needed for their development and operation are described in the architecture model, too.

The *operation model* defines processes, roles and tools for operation of the system and its components. There are many things to consider for the technical operation such as data backup, protection from failures etc. The operation model includes regulations regarding administration of the system as well as data supply.

The *organisation model* focuses on the form of organisation to implement the conceptual and technical tasks and requirements, i.e. the organizational structure (Grohmann & Stahl, 2010).

Besides the national SDI GDI-DE there are state SDIs for every federal state and sometimes even local SDIs on a more regional level like the GDI-Südhessen[12]. These SDIs are making use of reference models as well and thus selected ones are described in the next sections.

3.3.1.1 Reference Model of the GDI North Rhine-Westfalia

The reference model of the GDI North Rhine-Westfalia (GDI-NRW) is using sub models like the GDI-DE which are shown in figure 3.4. As can be seen in the figure it is using a *business* and *architecture model*, too. The operation model which was present in the GDI-DE is split up into three sub models in the GDI-NRW: the roles are described in the *role model*, the processes are defined in the *process model* and the development of technical components is laid down in the *implementation model*. The link between architecture, process and role model are scenarios which formalise processes. Figure 3.5 shows that the infrastructure is basically made of three layers: a client, services and storage.

The *business model* describes the information flow and the economic processes as well as the economic strategies and requirements which are brought to the GDI-NRW by participants. Furthermore it defines aspects of pricing and marketing as well as legal and organizational aspects.

The *architecture model* defines the components of the SDI (services and their interfaces, encodings, clients etc.) and their interrelations as well as relations of the components to national and international norms and standards (ISO TC 211/ISO 191XX series[13], OGC etc.). The architecture model is underlining the importance of services when building an SDI and mentions a variety of service types that should be included in the SDI, such as catalog, gazetteer, map, web pricing and ordering services.

The *role model* is concerned with the roles and their motives (*why* they want an SDI) and their expectations (*what* they want to do with the SDI). The role model specifies activities of the actors which can be public institutions, companies and private persons based on their role.

The *process model* specifies business processes which describe the relationships of actors (roles) inside the SDI. Processes are the base for services in an SDI because through the definition of information flows between actors (from user request until the finished "product" at the end, e.g. a certain map image with layer from different vendors) which are formalised into processes which in turn can be mapped to services. Multiple processes can be linked into process chains which use multiple services. The used services are depicted in figure 3.5 which also shows that UML has been used to model the infrastructure at one point.

[12] refer to SDI classification on page 19

[13] see subsection 2.2.1 on page 24

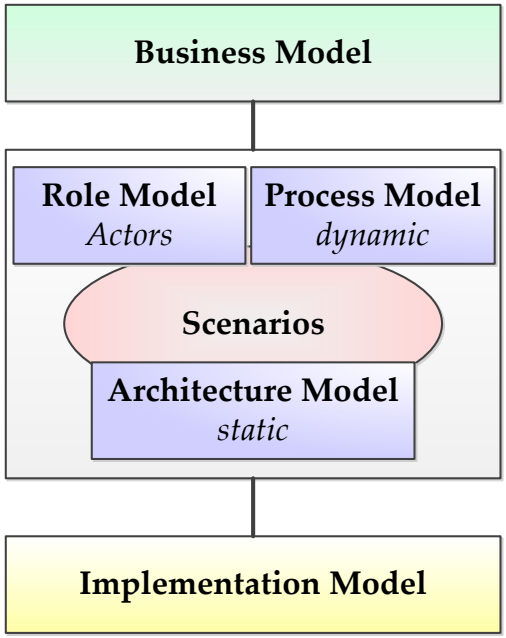


Figure 3.4: *Reference Model of the GDI North Rhine-Westfalia (modified after (Altmaier et al., 2003, p. 6))*

In the *implementation model* implementation specifications can be found which define aspects of the architecture model. These specifications list standards and contain information regarding technology decisions, development platforms, construction and reusability of components and performance (Altmaier et al., 2003).

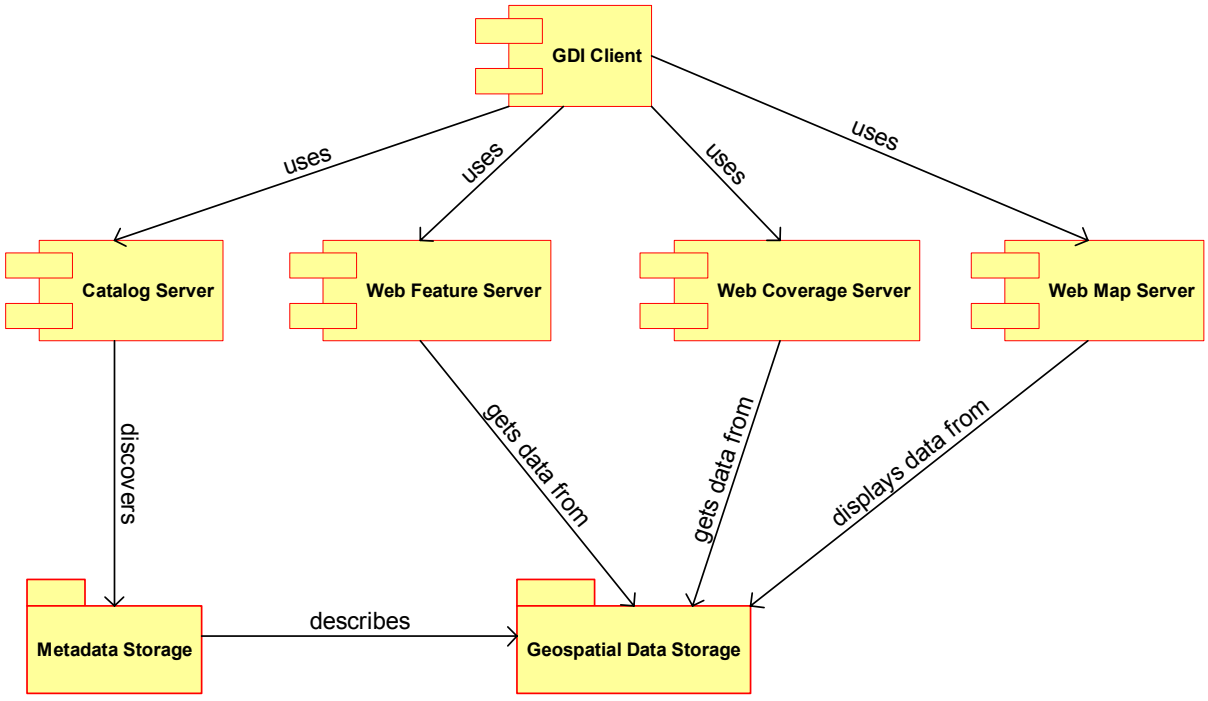


Figure 3.5: *UML component diagram of the GDI North Rhine-Westfalia (Brox et al., 2002)*

One point about the reference model of the GDI North Rhine-Westfalia which is relevant for reference models for SDIs in general is important to note: A reference model has to dynamically adapt technical and market oriented developments – it has to be developed further continuously.

3.3.1.2 Reference Model of the SDI Saxony

Another SDI inside the GDI-DE is the one of the federal state of Saxony. Its reference model portrays and describes the technical, organizational and political relations inside the Saxon SDI. The legal framework for the SDI Saxony is defined significantly through the INSPIRE directive. In contrast to GDI-NRW (see 3.3.1.1) the Saxon SDI is using fewer sub models as figure 3.6 shows. It uses even fewer sub models than the reference model of the GDI-DE. But these sub models encompass more or less everything the other reference models contain as the operating model encompasses a process and a role model for example. Figure 3.7 shows that the architecture consists of three layers, too. These three layers are equivalent to the ones of the reference model of North Rhine-Westfalia (client⇆application layer, services⇆service layer and storage⇆data layer) (GDI-Sachsen, 2009) and (Wytzisk, 2012).

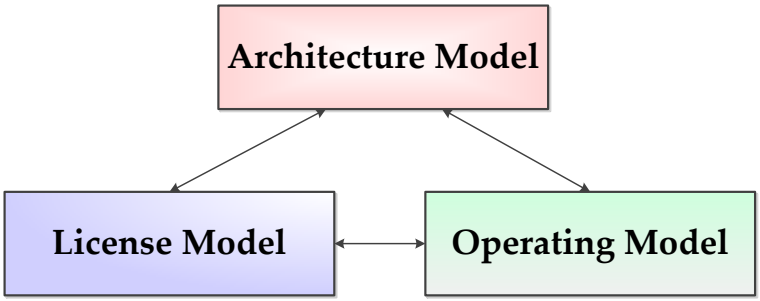


Figure 3.6: *Reference Model of SDI Saxony (modified after (Wytzisk, 2012, p. 9))*

The *operating model* describes the technical and organizational views on the SDI, the processes and the participating actors. That means that the operating model includes a process model which outlines tasks and actions needed for implementation and maintenance of the SDI by defining business processes as well as a role model which defines actors (people, institutions etc., called "subjects") and tools (called "objects") and their relationships inside the SDI.

Just like in the other reference models the *architecture model* describes the technical architecture required for the implementation (which means that the architecture model includes the considerations other reference models put into a separate implementation model) of the SDI. A part of the architecture model is the architecture concept which considers the architecture concept of the GDI-DE as well as requirements resulting from the INSPIRE directive. As depicted in figure 3.7 the architecture model is made of three layers where data flows from the bottom (data layer) to the top (application layer) through an intermediate layer – the services layer.

So far no reference model included a separate *license model*. This model occupies itself with legal considerations (including fees for acquisitions) and roles of providers and users and how they interact (Wytzisk, 2012) and [SDI Saxony website][14].

3.3.2 WRON Reference Model (WRON-RM)

Looking outside Germany and even Europe one can find Australia's Water Resources Observation Network (WRON) which was also built using a reference model which – in contrast to the reference models so far – is based on RM-ODP. WRON is about[15] *"[. . .] meeting Australia's current and future*

[14] http://www.gdi.sachsen.de/inhalt/konzept/refmod.html
[15] (O'Hagan et al., 2007, p. 1145)

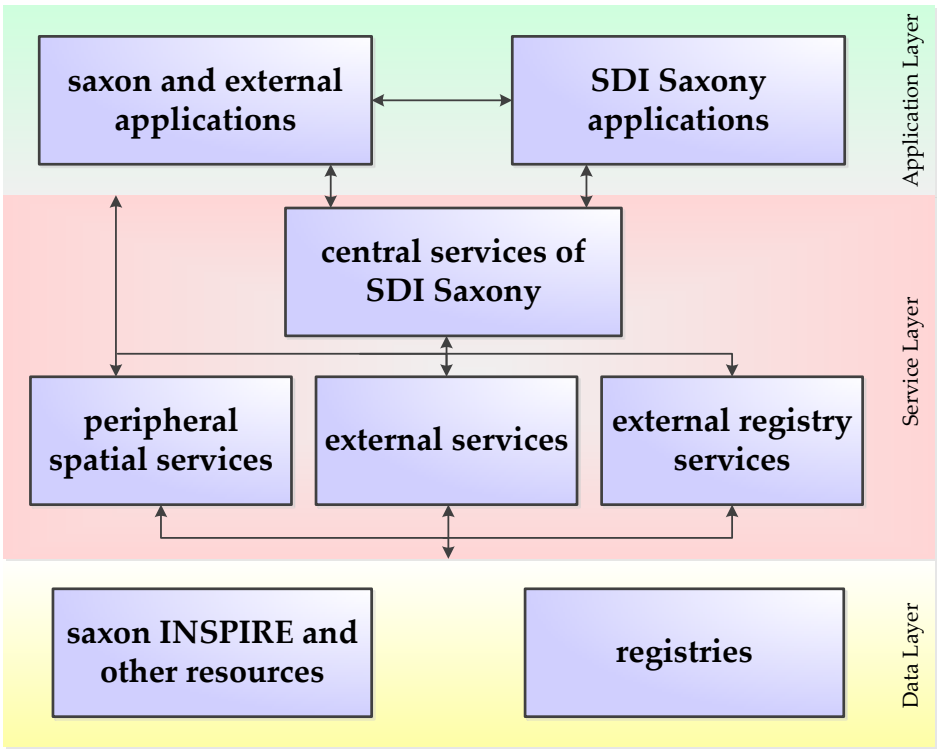


Figure 3.7: *Architecture of Saxon SDI (modified after (GDI-Sachsen, 2009, p. 4))*

water challenges [which] requires timely access to the current and forecast future status of water resources." WRON should couple data and services in order to improve forecasting and reporting. In detail WRON should[16]:

- link Australia's many water and water related data assets,
- harness new data streams from satellites and on-ground sensor networks and
- process and utilise water information.

The WRON Reference Model (WRON-RM) defines the architecture of the WRON and describes key information, government policy, use cases and standards used within the WRON. In order to build the architecture WRON-RM uses the structure of RM-ODP and thus uses its viewpoints (enterprise, information, computational, engineering, technology, see subsection 2.3.1) which are explained in the next paragraphs. It is important to note that as opposed to figure 2.9 on page 30 WRON-RM is using different relationships between the viewpoints than RM-ODP which is depicted in figure 3.8. Additionally WRON-RM developed a UML model of the architecture (O'Hagan et al., 2007, p. 1147).

Enterprise Viewpoint Describes the major architectural requirements, key business drivers, desired outcomes and adoption strategy of the WRON. It states the relationships to other relevant national and international initiatives such as the Australian Water Data Infrastructure Project (AWDIP) and National Data Network (NDN). In order to identify the constraints and obligations within the WRON (and which must be taken into consideration in all the other viewpoints) the enterprise viewpoint describes the environment within which the WRON will be developed.

Information Viewpoint Describes the information content of the WRON by identifying information elements, manipulations that may be performed on these elements and information flows.

Computational Viewpoint Describes the division of elements of the WRON into independent functional components by building a 'notional architecture'. These functional components are typically services and interfaces required for the construction of the WRON.

[16](O'Hagan et al., 2007, p. 1146)

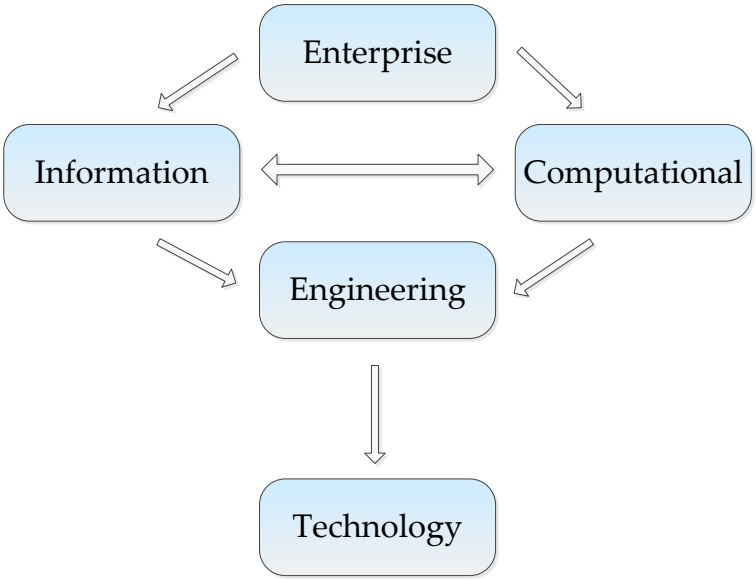


Figure 3.8: *Relationships between viewpoints in WRON-RM (modified after (Lemon et al., 2007, p. 5))*

Engineering Viewpoint Describes the outcomes from deployed components by providing a framework for evaluation of targeted performance, robustness, predictability of responses and processing requirements.

Technology Viewpoint Describes possible technical products computational, informational and enterprise structures as well as engineering mechanisms as independently as possible (Lemon et al., 2007, p. 4).

3.3.3 Digital Earth Reference Model (DERM)

After looking at Europe and Australia one can find the usage of a reference model even on a global level (on a regional respectively continental level there is the INSPIRE Generic Conceptual Model) with the Digital Earth Reference Model (DERM[17]) which also uses viewpoints like WRON-RM. *Digital Earth* was/is a vision verbalized by former US vice president Al Gore in 1998 (Gore, 1999) which aims at the provision of geodata about the earth freely over the internet (Craglia et al., 2012). Since this combines data from various sources it can be thought of as an SDI. This led to the foundation of a reference model for it – the Digital Earth Reference Model (DERM).

Although either not mentioning to be or being based on RM-ODP DERM is using similar viewpoints. As depicted in table 3.1 it uses the viewpoints *computation* and *information* on two different levels:

- **Implementation specifications** — *how* to implement requests, information portrayal etc.
- **Abstract models** — *what* is valid in principle (e.g. essential concepts, vocabulary etc.)

Table 3.1: *Viewpoints and levels of abstraction in DERM (Evans, 2001)*

	Computation Viewpoint *Service Invocation*	**Information Viewpoint** *Information Transfer*
Implementation specifications	Interface	Encoding
Abstract models	Behavior	Content

[17] http://cartome.org/draft-derm.htm

From a computational viewpoint implementation specifications define the interfaces that allow interoperability and from an information viewpoint they define the encoding of geodata so that it can be transferred between different systems. Abstract models define the expected behavior of the software or the whole system from a computation viewpoint. From an information viewpoint abstract models define the content of geospatial information (Evans, 2001).

DERM identified three layers and four major components the system should have. The layers *clients*, *middleware* and *servers* are shown in figure 3.9 along with the four main components of the system:

- **User applications** – Software ranging from analytical application to viewers which receive their data directly from the repositories or pre-processed from the middleware.
- **Geoprocessing services** – Process data (from simply drawing maps to advanced analytical functions) and provide them to the user or other services.
- **Content repositories** – Provide geodata from databases and so on as features, coverages, table etc.
- **Catalogs** – Allow metadata search to find geodata and services offering geodata.

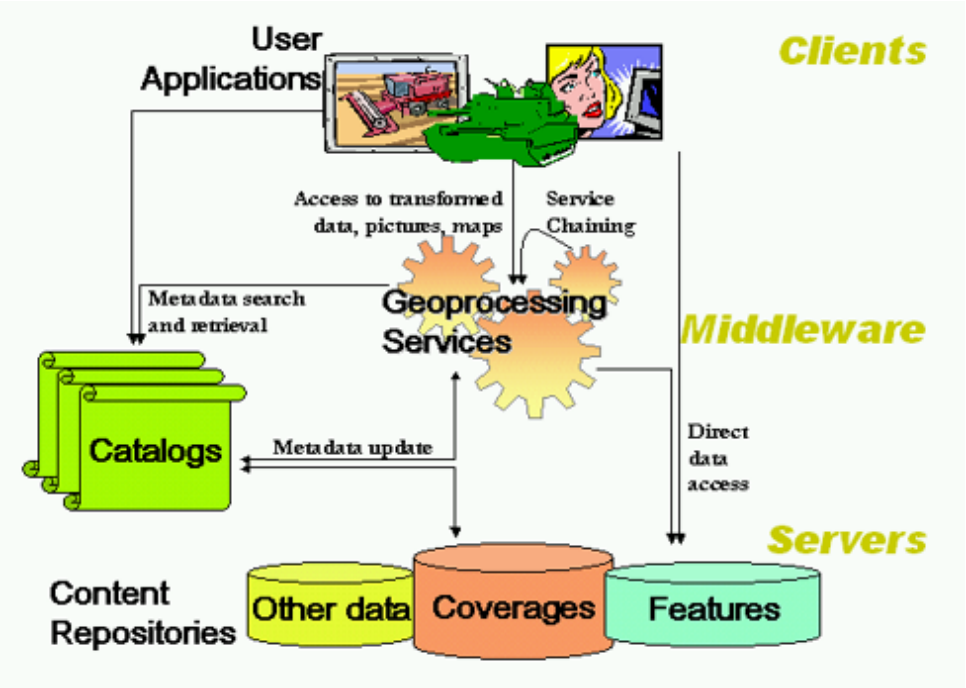


Figure 3.9: *Layers and compontents of the DERM (Evans, 2001)*

3.3.4 Conclusions

As figure 3.10 depicts the viewpoints and sub models of all reference models can be mapped to RM-ODP's viewpoints although it is especially hard for DERM because of its levels of abstraction. Thus RM-ODP seems to be a suitable base for the development of a reference model. Apart from that all the reference models looked at are using a layered architecture except for WRON-RM for which at least no information regarding this could be found. Hence a layered architecture seems to be a widely agreed on approach.

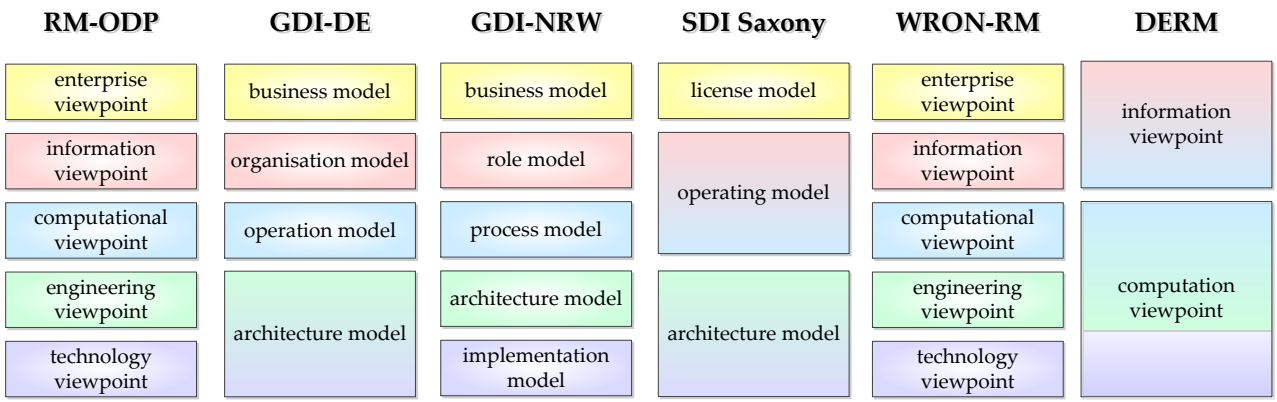


Figure 3.10: *RM-ODP's viewpoints and the corresponding viewpoints and sub models of selected reference models*

3.4 Existing marine vocabularies

Vocabularies describe the notional realm of a certain domain and are important because vocabularies ensure that all the terms are clear with respect to meaning and notation. Thus, vocabularies are important for the establishment of a marine SDI, too.

Notable examples for vocabularies in the environmental domain are the Global Change Master Directory (GCMD, http://gcmd.gsfc.nasa.gov) by the NASA with its keyword list[18] and the environmental thesaurus[19] (Umwelt-Thesaurus, UMTHES) by the Federal Environmental Agency (UBA). Another approach connected to the INSPIRE directive (see subsection 2.5.1 on page 42) is the GEneral Multilingual Environmental Thesaurus[20] (GEMET). All three include marine data respectively terms but are not marine-specific. However, the glossary by the Coastal and Hydraulics Laboratory[21] of the US Army Corps of Engineers is marine-specific because it *"[. . .] is a collection of terms that define physical, chemical, biological, environmental, ecological, geological and legal aspects of water and water resources."*

In Germany there are at least three word lists to be found which contain terms of the marine domain: the word lists "Küste" (subsection 3.4.1), NOKIS (subsection 3.4.2) and LHM (subsection 3.4.3).

3.4.1 Küste

The word list "Küste" by the German Coastal Engineering Research Council (KFKI) is very simple consisting of two columns in one Excel sheet only (see figure 3.11) and rather short, too, containing 167 terms. It just lists German terms in one column and their English translations in the other column if there is an English translation. But it also includes a very basic approach to semantics because it contains connections between related terms marked with "*s.*" (which stands for "siehe" (German) respectively "see" and which are also depicted in figure 3.11).

3.4.2 NOKIS

Another word list which handles relations between terms with a similar approach like "Küste" is the word list of the meta data base Nord-Ostsee-Küsten-Informationssystem (NOKIS[22]). Contrary to the "Küste" word list it sets the relations between terms for English translations only with the

[18] http://gcmd.gsfc.nasa.gov/learn/keyword_list.html
[19] http://data.uba.de/umt/
[20] http://www.eionet.europa.eu/gemet
[21] http://chl.erdc.usace.army.mil/glossary
[22] www.kfki.de/de/service/nokis

59	Vermessung	
	KFKI (Kuratorium für Forschung und	board for coastal engineering research
60	Küsteningenieurwesen')	(KFKI)
61	Knickboden	cracked soil
62	Konferenzen	conferences
63	Küstenabbruch s. Abrasion	
64	Küstenformen	coastal formations
65	Küstenforschung	coastal research
66	Küsteningenieurwesen	coastal engineering
67	Küstenschutz	shore protection
	Küstenschutzwerke	shore protection
68		structures

Figure 3.11: *Excerpt of word list "Küste"*

phrase "*s.*" (see figure 3.12). With almost 3000 terms in an Excel sheet it contains much more terms than the "Küste" word list. However, apart from relations between terms and translations it also contains descriptions respectively definitions of some terms in German or English (also depicted in figure 3.12). On top of that it contains references to where the translation or definition originated from with short labels like "*[12]*" or "*IW*". The meanings behind these short labels are broken down in a separate sheet.

	A	B	C	D	E	
1		[12]		Abbruchkante	escarpment	A more or less continuous line of cliffs or steep sl
2		[12]		Abbruchkante	scarp	see BEACH SCARP.
3		IW		Abdeckfundament	blanket bedding	see BLANKET FOUNDATION
4		IW		Abdeckfundament	blanket foundation	A layer or layers of graded fine stones underlying the natural bed material from being washed away.
5				Abfertigungszeit	handling time	
6				Abfertigungszeit	turnover time	
7	DIN	[1]	a) Allgemein: Unter dem Abfluss der Sc	Abfluss	runoff	
	DIN	[1]	Einjährige, nach Gesichtspunkten des Wasserhaushalts festgesetzte Zeitspanne	Abflussjahr	hydrological year	

Figure 3.12: *Excerpt of word list "NOKIS"*

3.4.3 LANIS Habitat Mare (LHM)

Relations between terms can also be approached through the use of hierarchies. This approach is used by the word list LANIS (Landschafts- und Naturschutzinformationssystem) Habitat Mare[23] (LHM) by the Federal Agency of Nature Conservation (BfN – Bundesamtes für Naturschutz). It has up to six hierarchy levels to group the terms and link them to each other (see figure 3.13). Although it does not contain definitions respectively descriptions of the terms it offers a variety of synonyms and translations (German and English) for each term (which are not depicted in the figure). These features and the fact that the list is split up into 20 separate Excel files containing all in all over 3000 terms (synonyms counted as one term) make LHM a technically mature thesaurus.

3.5 SKOS Tools

An important step towards the implementation of an interoperable infrastructure for marine data is to understand the terms used in the marine domain. For that purpose there are already word lists (mostly in Excel format). As 2.4.2.2 on page 37 pointed out the W3C specified SKOS which is a formal language to represent these word lists respectively controlled structured vocabularies. What is initially needed to make use of the existing word lists is a converter which transforms the existing word lists into the SKOS format. After achieving that a software program is needed to manage, edit

[23] www.lanis.de/habitatmare

	Hierarchieebene 0	Hierarchieebene 1	Hierarchieebene 2
1			
2			
3	LANIS Habitat Mare		
4		Marine Arten	
5			Tierarten
6			Pflanzenarten
7		Marine Biotoptypen	
8			Biotoptypen Deutschl
9			EUNIS Lebensraumty
10			FFH-Lebensraumtype
11			HELCOM Lebensrau

Figure 3.13: *Excerpt of word list "LHM"*

and make use of the terms (e.g. being able to use them when annotating data with metadata or for discovery purposes). The converted data has to be available for editors and other web services thus a web-based tool is needed. The W3C lists a variety of such tools relevant to SKOS on its Wiki page[24]. Based on this list subsection 3.5.1 looks at pre-existing conversion tools while subsection 3.5.2 deals with web-based SKOS management tools.

3.5.1 Conversion Tools

Apart from some tools converting *from* SKOS there are not many to be found which convert *to* SKOS in the above mentioned list by the W3C. One tool – *Skosify* which is discussed in 3.5.1.1 – was particularly interesting because it not only converts to SKOS but also improves the structure of a given vocabulary and validates it. OpenRefine (3.5.1.2) also detects inconsistencies. Other tools are outlined briefly in 3.5.1.3.

3.5.1.1 Skosify

According to its Google Code project page[25] *"Skosify is a tool that can be used to convert vocabularies expressed as RDFS [RDF Schema] and OWL [Web Ontology Language] into SKOS format. It can also be used to improve, enrich and validate existing SKOS vocabularies."* Skosify is written in Python and was developed in the FinnONTO project at the Semantic Computing Research Group (SeCo) in Finland and its code is open source.

Being able to convert from RDF Schema and OWL to SKOS Skosify also focuses on improving the quality of SKOS vocabularies. Skosify adjusts the structure of a given vocabulary using many processing steps of which selected ones are:

- Read input file and make sure the vocabulary has a `skos:ConceptScheme`
- If input was OWL transform classes/concepts, literals and relations according to the [types], [literals] and [relations] mappings defined in a configuration file (e.g. a mapping under [literals] could say `rdfs.label=skos:prefLabel` which means that all `rdfs.labels` will be converted into `skos:prefLabels`)
- Assure that all concepts have a `skos:inScheme` relation to a `skos:ConceptScheme`
- Assure that topmost concepts are pointed to through the usage of `skos:hasTopConcept` and `skos:topConceptOf` relationships
- Perform some validations (e.g. making sure that there is only one `skos:prefLabel` per language or detecting cycles in `skos:broader` relationships)

[24] http://www.w3.org/2001/sw/wiki/SKOS
[25] http://code.google.com/p/skosify/

3.5.1.2 OpenRefine

OpenRefine[26] (formerly known as *Google Refine*)

> *"[. . .] is a power tool for working with messy data, primarily for*
>
> *(1) detecting and fixing inconsistencies*
> *(2) transforming data from one structure or format to another*
> *(3) connecting names within your data to name registries (databases)"*[27]

To convert data from spread sheets and other sources the second feature is of interest. (Schandl, 2011) shows a way to use OpenRefine to transform spread sheets into SKOS format by using an RDF extension called *RDF Refine*[28] which allows export into RDF. According to (Maali et al., 2011, p. 4) the extension defines the structure of the RDF document through a skeleton. The skeleton enables the user to select the resources and literals and what they will become in the RDF document (e.g. a `skos:prefLabel` when the SKOS namespace is loaded in OpenRefine) and what their URIs will look like (depicted in figure 3.14). After these definitions the spread sheet can be exported into RDF/XML or Turtle format.

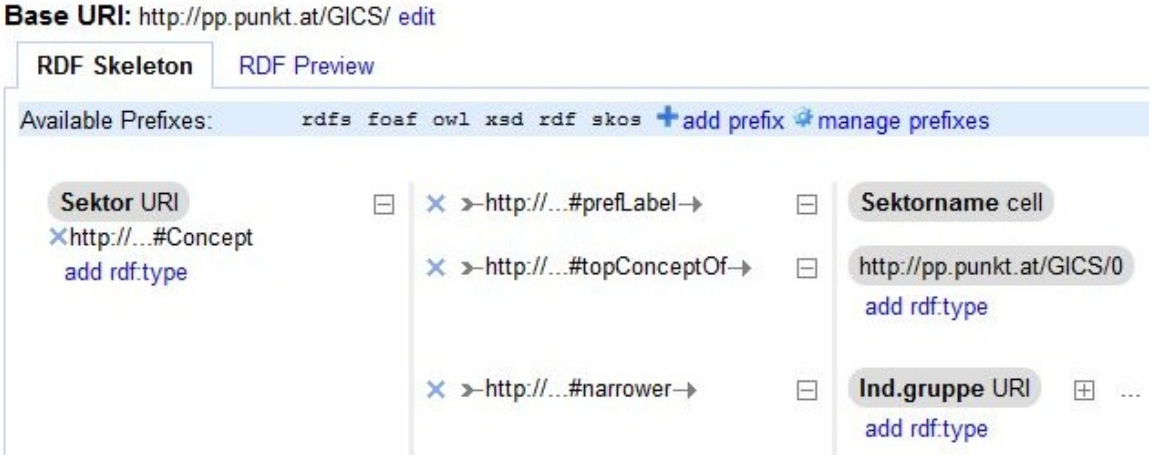


Figure 3.14: *Editing the RDF skeleton in OpenRefine (Schandl, 2011)*

3.5.1.3 Other conversion tools

A tool that promised *"[. . .] to convert OWL ontologies into SKOS terminologies"* according to its Google Code project page[29] is **OWLtoSKOS**. But after downloading (now called *skitter 0.1*) and deploying it it seems that this is not what the tool is about because when starting the Java Applet it gives a link which tells you to connect your Twitter account with an app called *ConceptCloud* – a development by the Rensselaer Polytechnic Institute (USA). The purpose of this app is to *"view the Twitter stream as it maps onto a SKOS terminology. See how concepts are mentioned, and how they link together."* What this means could not be evaluated because the tool (skitter 0.1) gave a NullPointerException after connecting the account to the ConceptCloud Twitter app.

Another tool written in Java is **OBO to SKOS Converter**[30] which as its name suggests converts Open Biological and Biomedical Ontologies (OBO) to SKOS (in various representations like RDF/XML, Turtle or OWL/XML) using the OWL API[31] and thus OWL as intermediate format. The tool converts

[26] http://openrefine.org/OpenRefine/index
[27] (Huynh, 2011, p. 2)
[28] http://refine.deri.ie/
[29] http://code.google.com/p/owltoskos/
[30] http://owl.cs.manchester.ac.uk/obotoskos/
[31] http://owlapi.sourceforge.net/

each OBO ontology into a SKOS concept scheme and every term from OBO is converted into a SKOS concept and a member of that concept scheme (http://www.cs.man.ac.uk/~sjupp/skos/index.html).

The **Zthes to SKOS Converter** which was developed for the thesaurus management software PoolParty (see 3.5.2.3) converts Zthes thesauri to SKOS (in RDF/XML format only) and is available as a service[32]. Zthes is an XML thesaurus format to define terms[33] which e.g. can be related to other terms using feature notes. These terms are converted to SKOS concepts with `zthes:termNotes` as `skos:scopeNotes` for example. To build URIs for the concepts the user can specify a base name space and a concept scheme ID which prepends the name space in the URI.

Another tool in the PoolParty realm is the **PoolParty Extractor**[34] which, however, is more focused on text mining and enriching an existing thesaurus.

AuScope – an Australian non-profit company which aims at facilitating the implementation of an infrastructure system for earth sciences – developed the **Excel to SKOS/RDF conversion tool** to process Excel sheets and convert their contents into SKOS[35]. The approach is based on *XL2XML* – a tool also developed by AuScope – which uses conversion rule sets. One of its rule sets generates a SKOS/RDF vocabulary file containing a set of concepts. Essentially the Excel sheets that can be processed have to be in a specific format: it has to contain certain fields and the data inside the table is a key-value list after all (SKOS property in one column and a value for it in the other column).

Voc2Skos[36] is part of the Marine Metadata Interoperability Project (MMI) Ontology Registry and Repository (ORR) and converts a given CSV file into SKOS in the upload process. The given CSV can simply contain the SKOS properties (e.g. `skos:prefLabel`) as column headings and an URI or ID column so that the concepts either get the URI stated in the column or the ID is appended to the URI given in the upload process or specified in the CSV. One can specify the URI by using the preamble of the CSV. The preamble supports several elements such as `ontologyURI` which specifies the URI for the ontology, the character used as `separator` and two elements which allow indentation, i.e. allow the use of hierarchies in more complex CSV files (`indent.string` and `indent.property`.

3.5.2 Web based Thesaurus Management Tools

To make the most of vocabularies a management tool is needed to easily

(1) add and change terms and
(2) use terms for metadata.

For (1) it is important to have an editorial system which allows it to lock terms which are changing. Roles need to be defined concerning authorization of whether to publish a change or a new term or leave it unchanged respectively do not add the new term. (2) requires management tools to make use of standardized web services to expose their knowledge to other applications e.g. in a metadata tool. Both functionalities of management tools of course require them to be web based.

When looking at the list of tools offered on the W3C SKOS Wiki page[37] which was mentioned at the beginning of this section only four could be found offering the functionalities just mentioned: TemaTres (3.5.2.1), MMI Ontology Registry and Repository (3.5.2.2), PoolParty (3.5.2.3) and iQvoc (3.5.2.4).

[32] http://demo.semantic-web.at:8080/SkosServices/zthes
[33] An example representation in Zthes format for the term *zoology* can be found here: http://zthes.z3950.org/schema/synapse.xml
[34] http://semanticweb.org/wiki/PoolParty_Extractor
[35] https://twiki.auscope.org/wiki/Grid/ExcelToRdfTool
[36] https://marinemetadata.org/mmiorrusrman/voc2skos
[37] http://www.w3.org/2001/sw/wiki/SKOS

3.5.2.1 TemaTres

According to its Wiki[38] *TemaTres "[. . .] is an open source vocabulary server [with a] web application for management [of] formal representations of knowledge, thesauri, taxonomies and multilingual vocabularies."*

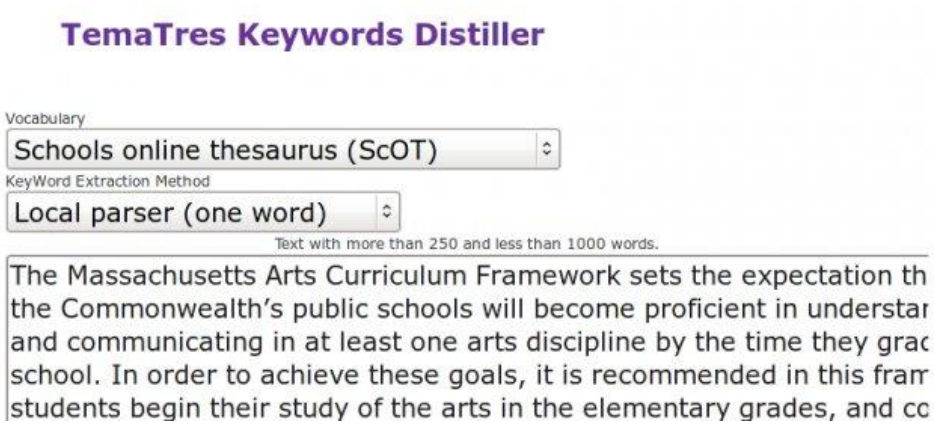


Figure 3.15: *Keyword extraction via TemaTres web services [SourceForge project page]*

The SourceForge project page[39] lists the requirements (PHP, HTTP Web server and Database server [e.g. MySQL]) and features of TemaTres of which some are:

- Expose vocabularies with web services (see figure 3.15)
- Support for multilingual thesaurus
- Relationship between terms (BT/NT, USE/UF, RT)
- Systematic or alphabetic navigation
- Export to SKOS/RDF, Zthes, TopicMaps, MADS, Dublin Core, VDEX, BS 8723, SiteMap, SQL, txt, WordPress XML
- Import thesauri from tabulated textfiles or data in SKOS format
- Scope notes, Historical and Bibliographical notes
- User management
- Workflow: candidate, accepted and rejected terms
- Allow to create user-defined relationships
- Search terms suggestion (*did you mean...?*)

3.5.2.2 MMI Ontology Registry and Repository

The Marine Metadata Interoperability Project (MMI) Ontology Registry and Repository (ORR) is a web-based *"[. . .] ontology service and ancillary tools [. . .] , that facilitate the creation and access of ontologies and mappings, and services that facilitate dynamic categorization of observations. [. . .] With the central goal of providing the marine community with supporting functions for semantic interoperability. [. . .] With this system, users and software agents can find ontology concepts and associated annotations using semantic web based query mechanisms."*[40]

[38] http://www.r020.com.ar/tematres/wiki/doku.php?id=:en:Inicio
[39] http://sourceforge.net/projects/tematres/
[40] (Rueda et al., 2009)

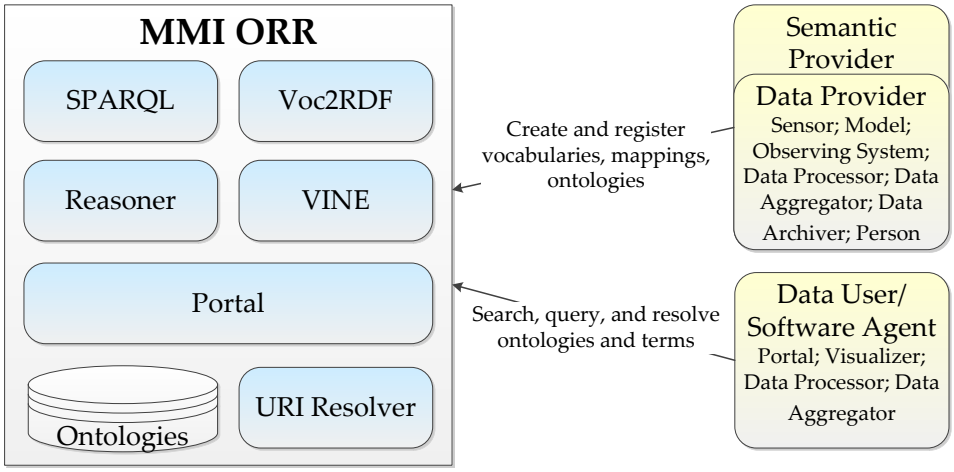


Figure 3.16: *Overview of MMI ORR components and interactions (modified after (Rueda et al., 2009))*

Figure 3.16 depicts the main components of MMI ORR which are implementing the main features of the system:

- Browse and edit ontologies and terms (portal and VINE)
- Create and register vocabularies, ontologies in standards-based format (RDF)
- Establish relations between terms to link vocabularies with others (mappings)
- Search for terms using free text or the query language SPARQL
- Provide web resolvable identifiers for vocabularies and terms (URI Resolver)
- Import vocabularies from CSV files (Voc2RDF[41])
- Make automatic inferences according to encoded rules in the underlying ontological model (Reasoner)

What above mentioned paper does not state – maybe because the features did not exist at the time – is that MMI ORR supports more formats to import[42] vocabularies or ontologies from: RDF/XML, Notation3 (N3), N-Triples and Turtle.

3.5.2.3 PoolParty

The first and only example of a proprietary and commercial tool is *PoolParty* which[43] *"[. . .] is a Thesaurus Management Tool (TMT) for the Semantic Web, which aims to support the creation and maintenance of thesauri by utilizing Linked Open Data (LOD), text-analysis and easy-to-use GUIs, so thesauri can be managed and utilized by domain experts without needing knowledge about the semantic web."*

Being written in Java and using AJAX (Asynchronous JavaScript and XML) PoolParty's main features include[44]:

- Create and maintain multilingual SKOS thesauri
- Publish thesauri with SPARQL or as Linked Open Data or API[45]
- PoolParty supports the import of thesauri in SKOS (in serializations including RDF/XML, N-Triples or Turtle) or Zthes format
- Browser based, collaborative thesaurus management[46]
- SKOS export (XML, Turtle, Trig and more)[47]

[41] now called Voc2SKOS, refer to https://marinemetadata.org/mmiorrusrman/voc2skos
[42] According to https://marinemetadata.org/mmiorrusrman/registerexisting
[43] (Schandl & Blumauer, 2010, p. 421)
[44] (Schandl & Blumauer, 2010, pp. 422)
[45] http://poolparty.biz/de/skos-without-sparql-poolparty-skos-api/
[46] http://poolparty.biz/poolparty-functionalities-features-at-a-glance/
[47] http://poolparty.biz/poolparty-functionalities-features-at-a-glance/

3.5.2.4 iQvoc

iQvoc which was developed by innoQ in conjunction with respectively for the Federal Environment Agency Germany[48] *"[. . .] is a new open source SKOS-XL[49] vocabulary management tool. [. . .] Its immediate purpose is maintaining and publishing reference vocabularies in the upcoming Linked Data cloud of environmental information, but it may be easily adapted to host any SKOS-XL compliant vocabulary."*

iQvoc is implemented based on a RESTful architecture as a Ruby on Rails application running on top of JRuby and uses the JavaScript library jQuery. For its implementation there were some key requirements given by the Agency which are also stated in above mentioned paper and are of course now features of iQvoc:

- SKOS-XL compliance
- Multilingualism
- Linked Data support
- Editorial team and workflow support
- Web interface for browsing and navigation
- Comfortable editing features with validation

According to the technical white paper (Bandholtz & Glaser, 2012) iQvoc offers a range of semantic services through HTTP GET requests which allow:

- Lookup of a concept, collection or label through its HTTP URI,
- Search for concepts, collections or labels by various attributes,
- Retrieval of sub-trees of the concept hierarchy,
- Getting a list of similar concepts with respect to a given term and
- Acquirement of a list of significant concepts that classify a given document.

3.6 Tools to evaluate performance and conformity of services

The SDIs introduced in section 2.1 on page 16 consist of several services and thus depend on their availability, performance and conformance to defined standards and regulations. Not available services are negative for other systems (e.g. portals) as well as frustrating for users. But even if services are available their performance is very important for satisfying usage. Furthermore the requests to and the responses of the services have to follow an agreed-on agenda to make sure that for example a users' GIS is able to understand the service and communicate with it. Hence conformity is another important factor worth consideration.

The approach presented focuses on the user expectations when using services which is why subsection 3.6.1 covers these expectations. The most important expectation of a user is that a service is available and delivers data fast. Thus subsection 3.6.3 is looking at tools testing and monitoring regarding their performance and availability. Another thing a user expects is interoperability which is the reason for subsection 3.6.2 looking at tools testing conformity. Because there is one tool – the GDI-DE testsuite – which is able to test both user expectations it is introduced in subsection 3.6.4 separately.

3.6.1 Quality of Service

Evaluation of services mainly falls into two categories: performance and availability on the one side and conformance on the other side. Service quality was defined by (Parasuraman et al., 1985, p 45) as the *"[. . .] discrepancy between [. . .] consumer expectations and the actual specifications established for a service."* This means that service quality shows what the user expected and what he/she perceived.

[48] (Bandholtz et al., 2010, p. 1)
[49] refer to 2.4.2.2 on page 39

Firstly a user expects a service to be available so that he is able to use it. An analysis from the year 2010 (Müller & Mandery, 2010, p. 814) where WMS found with Google were examined showed that 30 % of them were not available (responded with an error message although the request was formally correct and adjusted to the GetCapabilities document). Another quality characteristic of a service (as well as a SDI) is performance which is why this is an INSPIRE requirement (see subsection 2.5.1 an page 42). (Yang & Evans, 2008) coined the term Network GIS Performance (NGP) which *"[. . .] includes both the efficient use of Network GIS resources (CPU, memory, massive storage, geospatial data and geospatial services) and the perception of speed of a Network GIS."* INSPIRE defines performance briefly in (INSPIRE Consolidation Team, 2007, p. 17) as *"[. . .] how fast a service request can be completed."* From a user perspective the perception of speed of a service is what matters. (Yang & Evans, 2008) state that a study in 2001 showed that *"[. . .] the amount of time for a general user to wait for a web response without giving up is about 8 seconds"* (8-seconds rule). However, with widespread broadband connections this time might be lower nowadays. The analysis from the year 2010 cited above (Müller & Mandery, 2010, p. 814) examined the performance of spatial services, too. It showed that *"[. . .] the remaining 70 % had to deliver images (maps) which were 400 x 400 pixels and 86 % of the available services were able to deliver the images within less than one second."* When compared to the 8-seconds rule this is a very impressive outcome. However, performance measurement only makes sense when carried out over a longer time span. Hence permanence mentioned in (Müller & Mandery, 2010, p. 814) is important in conjunction with performance, too.

(Cibulka, 2013, pp. 31) lists several other attributes apart from availability which are affected by (bad) performance[50]:

- **Reliability*** – ability of a service to maintain a certain quality
- **Scalability** – ability of a system to increase the computing capacity according to current demands
- **Capacity*** – limit of the number of simultaneous requests with a guaranteed performance
- **Accuracy** – error rate generated by the web service
- **Accessibility*** – ability of a service to serve a client's request

(Cibulka, 2013, p. 32) (as well as (INSPIRE Consolidation Team, 2007, p. 20)) also lists interoperability as an attribute for quality of service. He states that *"[. . .] services should be interoperable [. . .] so that developers [. . .] do not have to think about which programming language or operating system the services are hosted on."*

To be interoperable services have to conform to standards and specifications. Conformance is the second main area for service evaluation. (Hogrebe, 2012, p. 3) points out that there is a strong relationship between conformance (called *regulatory* by INSPIRE in (INSPIRE Consolidation Team, 2007, p. 19)) and interoperability because from conformance derives interoperability. If a service fulfils the requirements of standards respectively specifications it will be able to operate together with other services also complying with these standards respectively specifications.

3.6.2 Tools concerning conformity

Conformity asks if a service, its data and metadata is[51] *"[. . .] in conformance with the rules, the law, complian[t] with standards, and the established service level agreement."* For a download service to be compliant to INSPIRE requirements it has to support a `language` parameter in its Get-Service-Metadata response and its response time for Get-Download-Service-Metadata requests has to be ten seconds at most for instance (INSPIRE Network Services Drafting Team, 2009, p. 12 and 14). Tools testing INSPIRE conformity will be introduced in 3.6.2.2 and 3.6.2.3. However, initially the service should be compliant to OGC specifications so that it understands the requests INSPIRE presupposes. That is why a tool by the OGC testing its services is presented firstly in the next section.

[50] Marked with an asterisk are attributes used as minimum performance criteria for the INSPIRE Network Services (INSPIRE Consolidation Team, 2007, p. 17).

[51] (INSPIRE Consolidation Team, 2007, p. 19)

3.6.2.1 OGC TEAM Engine

The Open Geospatial Consortium (OGC) pursues an initiative to verify specification compliance with its standards (Compliance and Interoperability Testing Initiative, CITE, as part of their OGC Compliance Testing Program). The tests are available online[52] and are based on the open source Test, Evaluation, And Measurement (TEAM) Engine[53] which[54] *"[. . .] executes test suites written using the OGC CTL [Compliance Test Language] test grammar or the TestNG framework."* The TEAM Engine is written in Java and can be used to execute test suites by command-line, web interfaces and a REST API (for TestNG-based suites). 5.4.1.1 on page 112 will detail some of its featured tests. At the time of writing the online version of the TEAM engine provided by the OGC offers test suites for:

- Catalog Service-Web (CSW) 2.0.1/2.02.
- Sensor Observation Service (SOS) 1.0.0
- Sensor Planning Service (SPS) 1.0
- Web Coverage Service (WCS) 1.0.0/1.1.1
- Web Feature Service (WFS) 1.0.0/1.1.0
- Web Map Context Documents (WMC) 1.1.0
- Web Map Service (WMS) 1.1.1/1.3.0
- Web Registry Service (WRS) 1.0

3.6.2.2 Neogeo WMS INSPIRE tester

In its scope the online *WMS INSPIRE tester*[55] by Neogeo Technologies states that *"[. . .] this application checks if the WMS GetCapabilities response is compliant to the requirements and recommendations of the INSPIRE view services technical guidances 3.0."* The requirements and recommendations[56] are numbered and the application makes use of this enumeration in the sense that it presents the test results to the user referring to requirements with `TG_Req#XX` where TG means *TechnicalGuidance*, Req means *Requirement* and the number after the hash sign (in this case XX) indicates the number of the requirement in the INSPIRE document. Recommendations follow the same scheme with `TG_Rec#XX`. Furthermore the scope states that the application *"[. . .] does not test the service against the requirements of the WMS specifications. It focuses on the INSPIRE specificities. At the moment, the test suite is not complete (by far)."*

Figure 3.17 shows an excerpt of the test results for a WMS handling MSFD related data by the Wadden Sea National Park Administration of Lower Saxony (NLPV). Although displayed just as *errors* all the errors visible in the figure are non-critical. However, there is no explanation of when an error is critical or when it is not which is not surprising due to a total lack of documentation or other information about the application. The only critical error discovered was *"INSPIRE view services shall implement WMS 1.3.0 standard (ISO 19128). This service does not implement this version."* when given a WMS 1.1.1 as input. Further common non-critical errors not depicted in the figure are inter alia *"BoundingBox missing for some CRS - TG_Req#36"* and *"The element <inspire_common:MetadataUrl> is not present in the element <inspire_vs:ExtendedCapabilities> - TG_Req#06"*.

3.6.2.3 INSPIRE Geoportal Metadata Validator

In its *about* section the INSPIRE Geoportal Metadata Validator[57] (European Commission, 2013) by the European Commission defines its purpose and scope as *"[. . .] to test the compliance of metadata of spatial*

[52] http://cite.opengeospatial.org/teamengine/
[53] http://sourceforge.net/projects/teamengine/
[54] (OGC, 2013, p. 1)
[55] http://inspire_tester.neogeo-online.net
[56] http://inspire.jrc.ec.europa.eu/documents/Network_Services/TechnicalGuidance_ViewServices_v3.0.pdf
[57] http://inspire-geoportal.ec.europa.eu/validator/

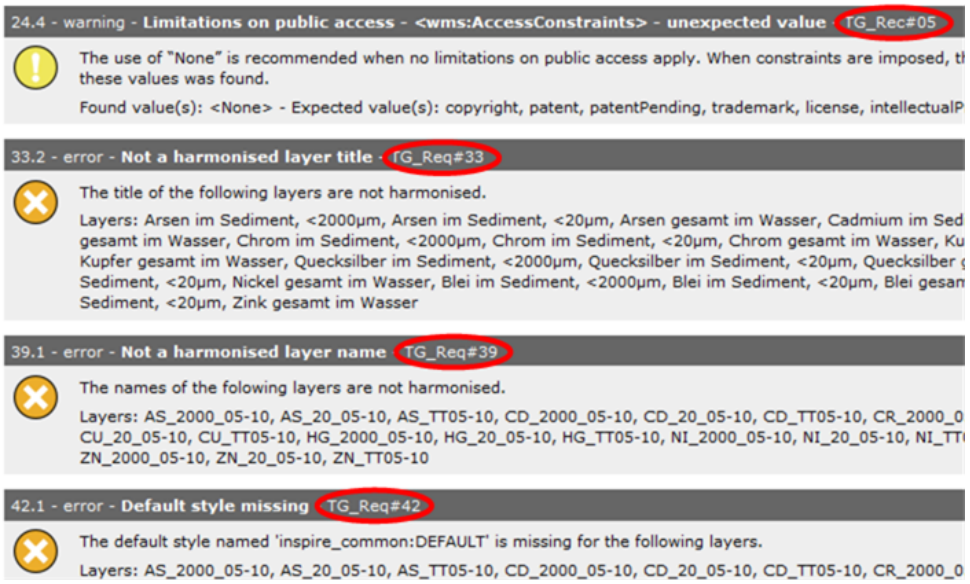


Figure 3.17: *Neogeo WMS INSPIRE tester test results*

data sets and services with the Metadata Technical Guidance." This indicates that while the Neogeo WMS INSPIRE tester (see 3.6.2.2) focuses on INSPIRE view services only the official INSPIRE Geoportal Metadata Validator offers tests for that and additionally for

- Metadata encoded in EN ISO 19139,
- Discovery Services (OGC CSW 2.0.2 AP ISO 1.0.1 with INSPIRE Extensions) and
- Download Services (ATOM, ISO 19142 (OGC WFS 2.0.0) with INSPIRE Extensions)

on top of that. For INSPIRE view services (Web Map Services) common validation issues inter alia include

- *the INSPIRE Extended Capabilities element could not be read,*
- *the metadata element "Resource Type" is missing, empty or incomplete but it is required,*
- *the metadata element "Mandatory Keyword" is missing, empty or incomplete but it is required,*
- *the metadata element "Layers" is missing, empty or incomplete but it is required* and
- *the metadata element "Conditions For Access And Use" is missing, empty or incomplete but it is required.*

3.6.3 Tools testing performance and availability

Subsection 3.6.1 introduced inter alia performance and availability as characteristics for quality of service. (Cibulka, 2013, p. 1) is pointing out that these characteristics *"[. . .] are dependent on the number of users, the type of operation [GetMap, GetCapabilities, GetFeature etc.] which is being performed, the volume of the data being processed, and a number of other parameters (hardware, software, network, etc.)."* However, the number of users is depending on the time of day, as (Drerup, 2010, p. 17) is stating. The users are producing load on a server whereas the more users the higher the load. Unfortunately none of the tools presented in the next sections are able to produce load artificially. This is the reason why 3.6.3.4 is looking at further tools handling this issue. Furthermore (Cibulka, 2013, p. 1) discovered that performance depends on the map scale as well as the location. With some effort it would be possible to reflect this behaviour of spatial services with the serviceMonitor presented in 3.6.3.2 and with the GDI-DE Testsuite discussed in subsection 3.6.4. Both tools allow to define multiple tests on one service with different requests which means that different scales and locations can be used to test the service. MapMatters (3.6.3.1) and the Service Status Checker (3.6.3.3) are not able to do this because requests cannot be defined by the user.

All the tools presented also test availability because it is the necessary requirement for performance testing. However, there are tools on the market which suggest that they test availability *and* performance but which just test availability. The Geomonitor by GeoInfoMarkt.org is such an example. It states on its website[58] that it offers statistics for availability *and* performance of services. But when you look at a service[59] it only shows the availability in the last 3, 6 and 9 hours.

3.6.3.1 geOps MapMatters

MapMatters[60] is a portal developed by geOps which main purpose is to catalogue web map services (WMS) and offer text search as well as search by geographical extent to users who want to find WMS. MapMatters uses both robots and user input to fill its database of WMS. Furthermore MapMatters offers a monitoring component which allows to record availability and response times as figure 3.18 shows. This is a feature a user has to activate and when he does requests for every layer of the services are sent half-hourly for one month. However, validity of layers respectively services is a thing MapMatters examines through its XML validation of the GetCapabilities document and errors messages like `Invalid BBOX` for layers, too (Müller & Mandery, 2010, pp. 814).

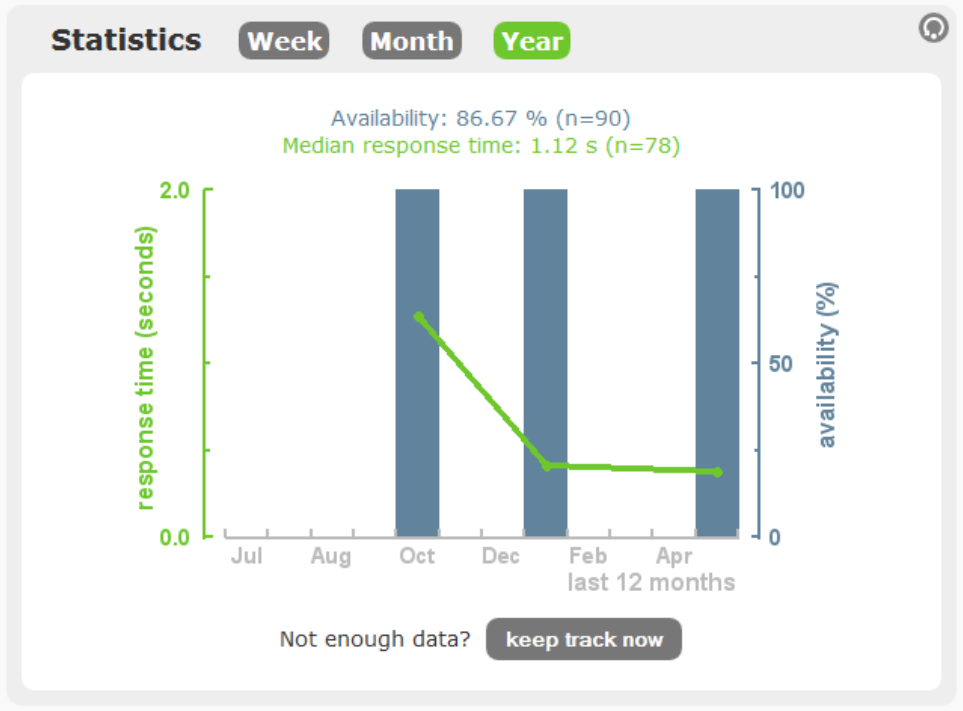


Figure 3.18: *MapMatters monitoring results of a WMS*

3.6.3.2 sdi.suite serviceMonitor

The serviceMonitor by con terra monitors different service types such as OGC WMS, WFS, WCS as well as ArcIMS and ArcGIS and INSPIRE Network Services. The time intervals as well as the thresholds regarding availability and performance which should not be underrun can be configured and if a service falls below the prescribed quality requirements personnel can be informed via E-Mail or SMS. For INSPIRE Network-Services the INSPIRE Quality of Service parameters are already preconfigured. Additionally figures about availability and performance of the services can

[58] http://geoinfomarkt.org/modules/geomonitor/
[59] http://geoinfomarkt.org/modules/geomonitor/info.php?service=11
[60] http://www.mapmatters.org

be displayed as reports with diagrams (see figure 3.19) or statistics which can be exported to MS Excel format. The serviceMonitor is integrated in con terra's sdi.suite but can be used with other products, too (con terra, 2013).

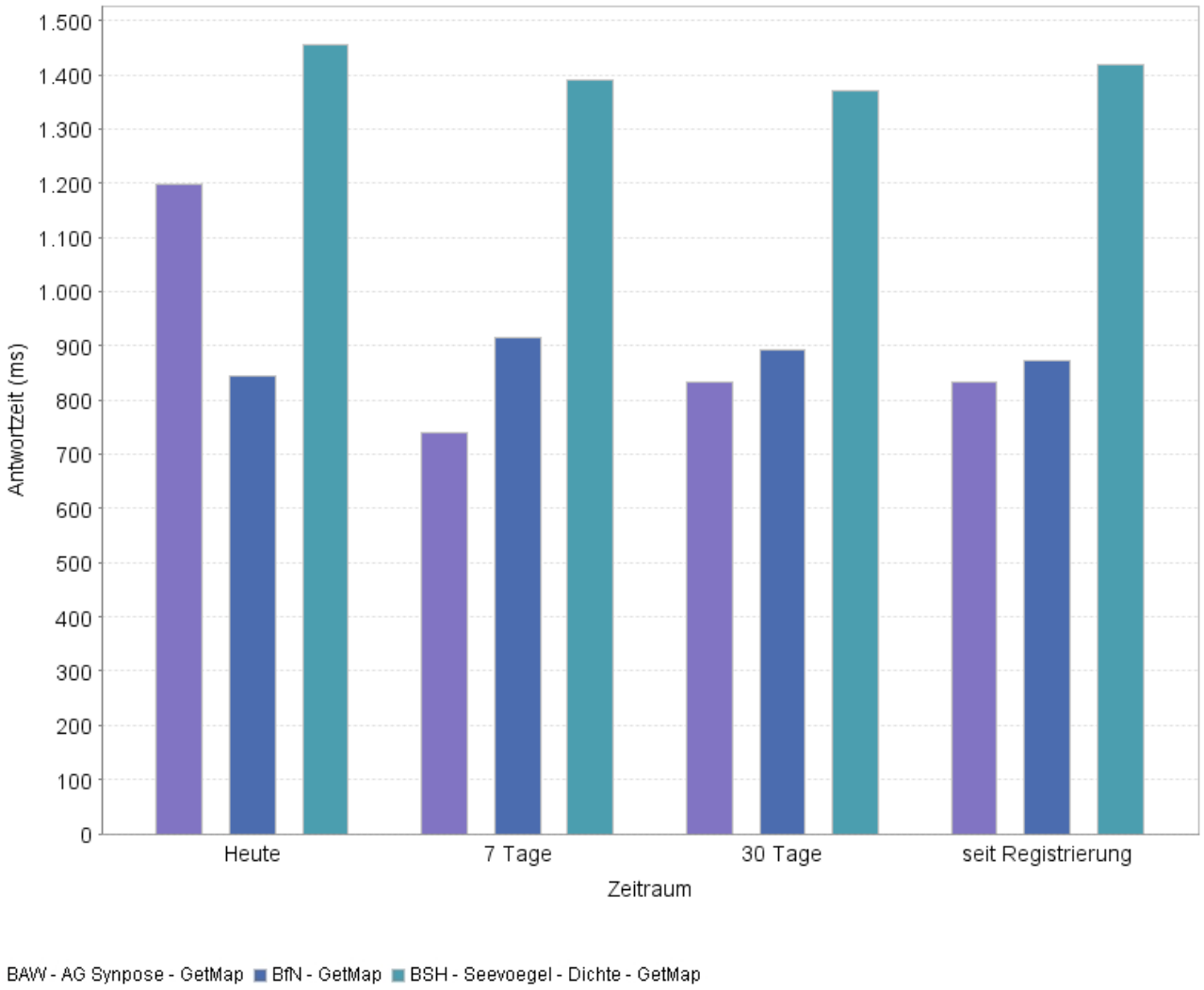


Figure 3.19: *serviceMonitor results regarding response times for three services displayed as a diagram*

3.6.3.3 FGDC Service Status Checker (SSC)

The objectives of the Service Status Checker (SSC) provided by the US American Federal Geographic Data Committee (FGDC) are summed up by (Anthony & Nebert, 2012, p. 3) as

- Monitor and score spatial web services,
- Notify owners of service issues,
- Spot performance issues,
- Determine why something is broken and
- Share your service testing results with others.

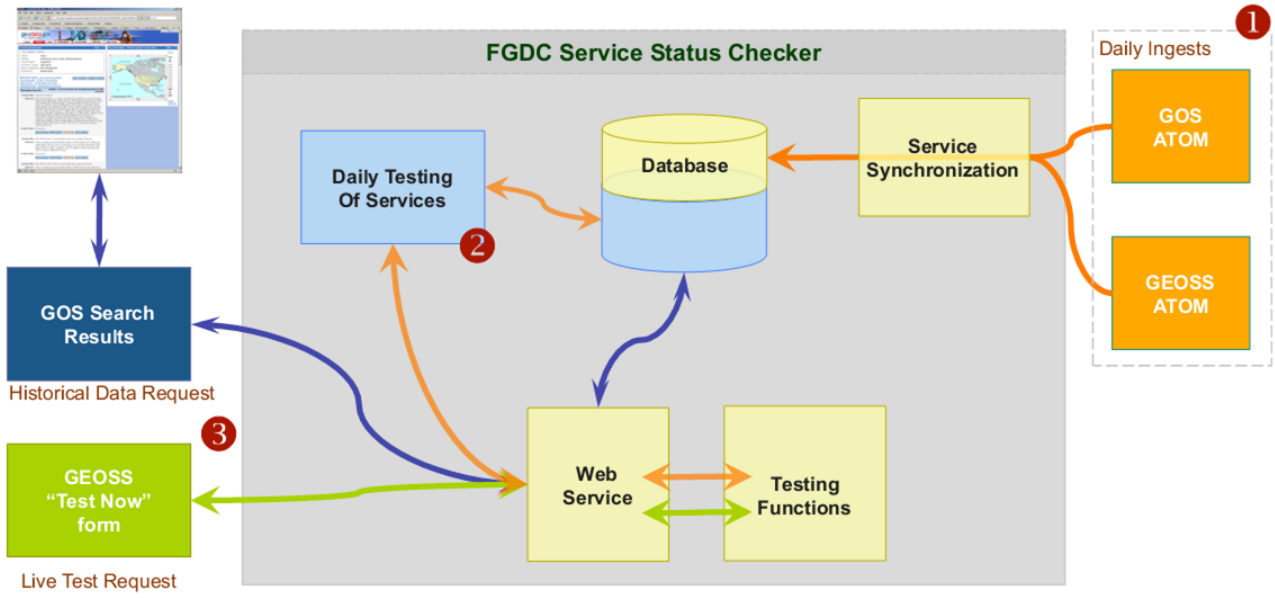


Figure 3.20: *SSC overview (Nebert, 2009, p 20)*

```
1  <feed xmlns="http://www.w3.org/2005/Atom" xmlns:gos="http://www.geodata.gov/gos_atom"
       >
2    <title>MDI-DE Services</title>
3    <updated>2013-06-04 00:00:00</updated>
4
5    <entry>
6      <id>WMS_BSH_CONTIS_Facilities</id>
7      <title>WMS BSH CONTIS - Facilities</title>
8      <updated>2013-06-04 15:00:00</updated>
9      <gos:serviceType>wms</gos:serviceType>
10     <gos:serviceUrl>http://gdisrv.bsh.de/arcgis/services/CONTIS/Facilities/MapServer/
           WMSServer</gos:serviceUrl>
11     <gos:esriServiceName></gos:esriServiceName>
12   </entry>
13
14   ...
15 </feed>
```

Listing 3.1: *ATOM feed for MDI-DE services needed by SSC (excerpt)*

The SSC offers two main access methods. The first method allows to look at how services performed in the past and thus is called *Historical Tests*. Here lists of services are provided by catalogs or a registered user via ATOM feeds (see listing 3.1 and (1) in figure 3.20) and can then be batch tested at a specific time interval (two times a day) with the results archived in a database (2). The results are available through a web service (through an API using GET requests with output in JSON format) and through detailed HTML service reports (3). The other access method is called *Live Tests* where real-time tests are performed through web service requests (REST API with GET and POST requests) on given services and where the results are returned to the user immediately (see figure 3.21) (Nebert, 2009, p. 19) and (Sotis, 2009, p. 70). The API requests return the date and time of test results, service type and service ID and a test results summary as well as the detailed test results[61]. Apart from this API SSC offers a developer API which[62] *"[. . .] allows you to programmatically access the service test/score*

[61] http://registry.fgdc.gov/statuschecker/documentation.php#rest-api
[62] http://registry.fgdc.gov/statuschecker/account/api-request.php – log-in required

results for use in your own applications."

According to its documentation[63] SSC supports the following service types:

- Data Services
  - WMS
  - WFS
  - WCS
  - SOS
  - ESRI ArcIMS Image
  - ArcGIS Map Server
  - ArcGIS Image Server
  - ArcGIS Feature Server
- Metadata Services
  - Z39.50
  - Web Accessible Folder (WAF)[64]
  - CSW

Figure 3.21: *SSC "Service Report Card" for the BSH service CONTIS/Facilities (excerpt)*

For the evaluation of availability and performance of the CGDI (Canadian Geospatial Data Infrastructure) (Gao et al., 2010, p. 3) discovered a problem using SSC stating that the *"[. . .] response time provided by SSC was not completely satisfactory, because SSC used the entire world range to request maps from geospatial Web Services in testing."* This led to the decision to use web load testing tools (see 3.6.3.4) to evaluate *"[. . .] time latency, response speed, performance testing, and load testing metrics"*.

3.6.3.4 General web service (load) test tools

As already stated performance (defined in inter alia (INSPIRE Consolidation Team, 2007, p. 17)) *"[. . .] represents how fast a service request can be completed."* This response time can be measured with all the test tools presented so far. However, (INSPIRE Consolidation Team, 2007) furthermore defines capacity which *"[. . .] is the limit of the number of simultaneous requests which should be provided with guaranteed performance."* This means that there is a strong correlation between the performance and the amount of simultaneous users, i.e. requests. Unfortunately none of the tools presented so far are able to simulate a certain number of simultaneous requests. This is the reason why this section looks at such tools.

The website SoftwareQATest.com by Rick Hower studies quality assurance and testing since 1996. In *Web Site Test Tools and Site Management Tools* it lists 66 tools in the category *Load and Performance Test Tools* of which Proxy Sniffer, WAPT, Apache JMeter and ApacheBench will be discussed briefly. These were selected because they were already used for testing spatial web services.

[63] http://registry.fgdc.gov/statuschecker/documentation.php#service-types
[64] https://geo-ide.noaa.gov/wiki/index.php?title=Web_Accessible_Folder

Proxy Sniffer – formerly by Engineering Office David Fischer now Apica – is a load and stress testing tool able to record web sessions and automatically create Java test programs from the recording (Hower, 2013). Furthermore according to the product website[65] reports with charts and tables can be output as HTML or PDF documents and more than one million simultaneous users sending requests can be simulated. However, testing OGC services has to be done using other tools[66] *"[. . .] with its proxy destination to Proxy Sniffer."*

WAPT (Web Application Load, Stress and Performance Testing[67]) by SoftLogica LLC and according to (Horák et al., 2009, pp. 11) it *"[. . .] offers transparent design of scenarios, settings of virtual users, number of iterations, interval between runs and delay between users."* (Horák et al., 2009, pp. 11) also points out that outputs are presented as reports and graphs and furthermore applies the tool to OGC services. However, (Drerup, 2010, p. 14) states that WAPT is not useful to meet INSPIRE requirements because it is *"not able to measure the single download time in a response. It could only show the overall response time which includes also the time the server needs to process a request."* INSPIRE requires each duration to be measured separately.

Apache JMeter is a Java desktop application from the Apache Software Foundation dating back to 2001[68]. Thus it is a very stable and reliable tool to test functional behaviour, measure performance and analyze it graphically. It can test both static and dynamic resources and is able to simulate load on a server, network or object (Hower, 2013). In contrast to WAPT JMeter[69] *"[. . .] can measure not only the overall load time of a response, but also the latency, the time when the first byte of the response is received."* These features make it suitable to test services against INSPIRE requirements.

ApacheBench comes with the standard installation of Apache's HTTP server. It is a command-line tool to benchmark the (Apache) HTTP server installation. According to its documentation[70] it *"[. . .] especially shows you how many requests per second your Apache installation is capable of serving."*

3.6.4 GDI-DE Testsuite

The GDI-DE is the national SDI of Germany and one of its components is the GDI-DE Testsuite which was developed by the Coordination Office SDI at the Federal Agency for Cartography and Geodesy (BKG) in accordance with the Steering Commitee GDI-DE between end of 2010 and its public release in mid 2011. The purpose of the Testsuite according to (Hogrebe, 2011, p. 1) *"[. . .] is to provide a system to test the conformity of services and datasets with the requirements of INSPIRE and SDI Germany (GDI-DE)"* and in addition to *"[. . .] facilitate and automate the monitoring required by INSPIRE."* The Testsuite can be used as an online web application[71], accessed via its API or installed and used locally. (Hogrebe, 2012, p. 4) summarises the objectives of the GDI-DE Testsuite:

- Support interoperability in GDI-DE (common tool)
- Support the implementation process of GDI-DE and INSPIRE (common understanding of the relevant standards and specifications)
- Support INSPIRE Monitoring (conformity indicators)

[65] http://www.proxy-sniffer.com
[66] (Gao et al., 2010, p. 3)
[67] http://www.loadtestingtool.com/index.shtml
[68] http://archive.apache.org/dist/jakarta/jmeter/old/release/v1.0/
[69] (Drerup, 2010, p. 14)
[70] http://httpd.apache.org/docs/2.2/programs/ab.html
[71] http://testsuite.gdi-de.org

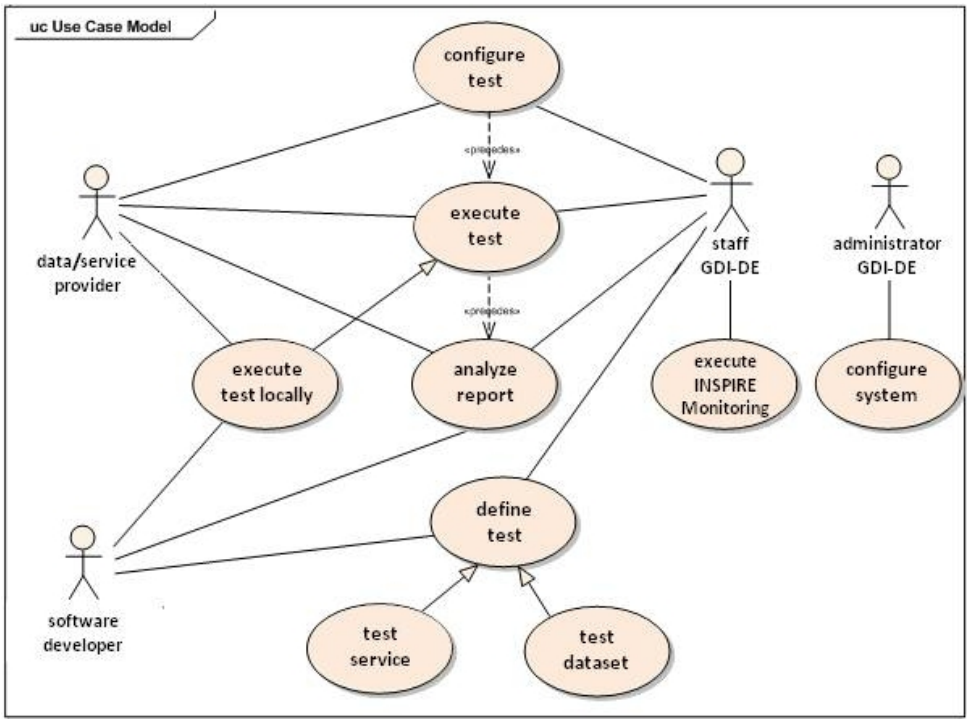


Figure 3.22: *Use Case Model GDI-DE Testsuite (Hogrebe, 2011, p. 2)*

Figure 3.22 illustrates the main stakeholders and use cases of the GDI-DE Testsuite which is based on OGC's TEAM Engine and uses its Compliance Test Language (CTL) (Hogrebe, 2012, p. 6). The annex of the user manual of the Testsuite (Böhme & Hogrebe, 2012, pp. 25–27) lists all available test classes with three main application domains:

- Metadata
  - GDI-DE – conformity to ISO 19115/19119 and optionally GDI-DE conventions
  - INSPIRE – conformity to INSPIRE metadata requirements
- Catalogue Services
  - OGC CSW 2.0.2 AP ISO 1.0 – conformity to OGC Catalogue Services Specification 2.0.2 - ISO Metadata Application Profile 1.0
  - INSPIRE Discovery Service 3.0 – conformity to INSPIRE Discovery Service requirements
- Map Services
  - GDI-DE WMS 1.1.1/1.3.0 – conformity to OGC Web Map Server Implementation Specification 1.1.1/1.3.0
  - INSPIRE View Service based on WMS 1.1.1/1.3.0 – conformity to INSPIRE View Service requirements with optional INSPIRE Quality of service tests (availability and performance, see figure 3.23)

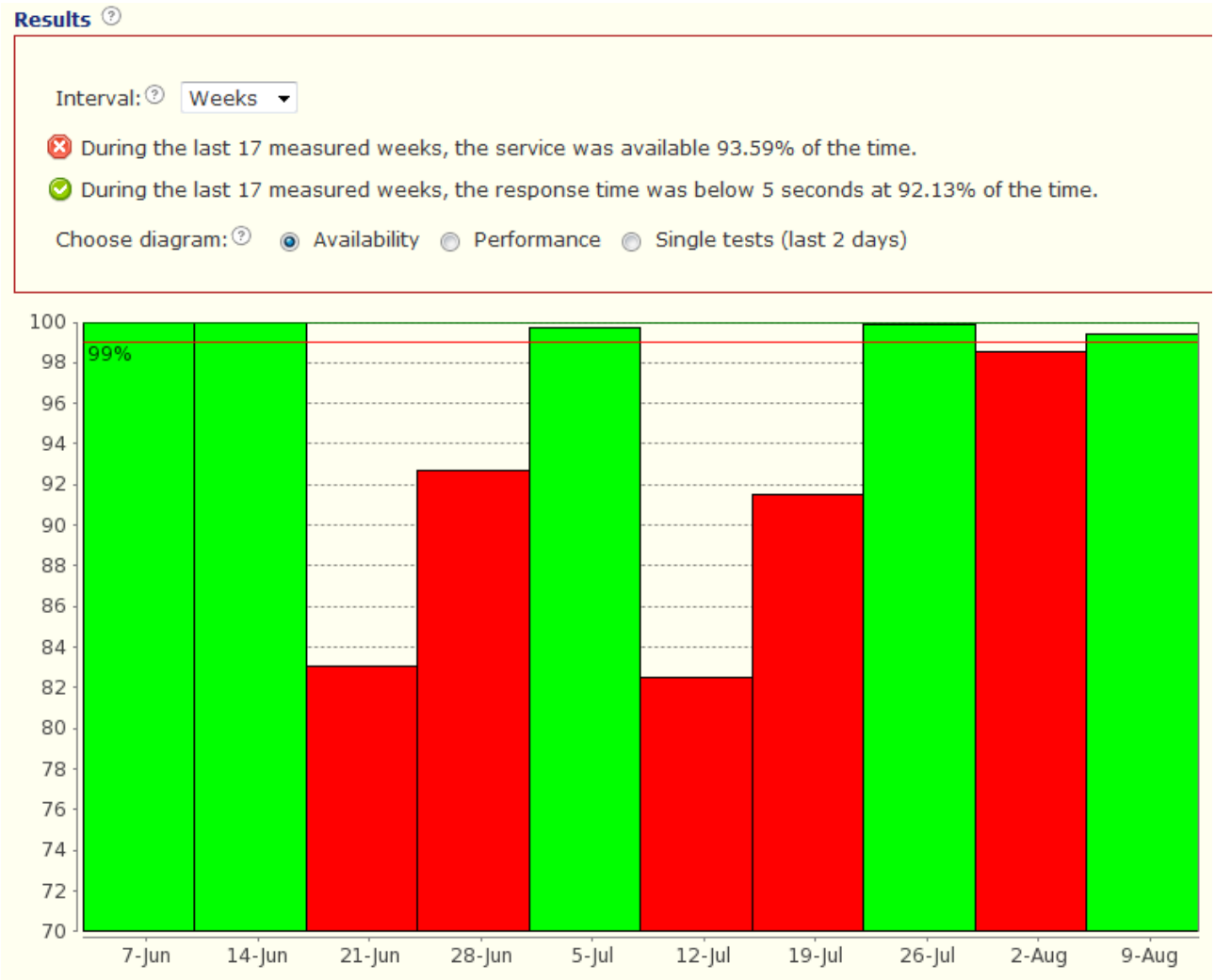


Figure 3.23: *Results for quality of service of BSH service SeaSurfaceTemperature*

4 Evaluation of existing MSDIs

> *"Evaluation is about finding answers to questions such as 'are we doing the right thing' and 'are we doing things right'."*
>
> (Steudler et al., 2008, p. 194)

In order to assess marine SDIs (MSDIs), a base is needed in order to compare them from a neutral point of view. That is why (in section 4.1) an evaluation framework for MSDIs is built first. This framework is based upon the components of SDIs and is composed of existing frameworks to evaluate SDIs. However, only aspects are included that apply to MSDIs and correspond to SDI components. Furthermore, the framework is expanded to meet the requirements of marine SDI evaluation. The last step is to apply SMART criteria on the indicators of the framework built (Rüh & Bill, 2012b). After that this framework is used to assess various marine SDIs around the globe (see section 4.2) (Rüh et al., 2012a).

4.1 Building an evaluation framework

To build an evaluation framework for MSDIs existing spatial data infrastructure assessment approaches were used as bases and were expanded to meet the requirements of the marine domain (see subsection 4.1.1). Based on researching literature in this field the components useful for marine SDI evaluation are selected and are then augmented by those needed especially in the marine context. The resulting indicators are merged in subsection 4.1.2 and are then assessed in subsection 4.1.4 to verify their fitness for use in an evaluation. But before that all the indicators are described in subsection 4.1.3. Through the assessment a number of indicators were identified for being considered in the framework for MSDI evaluation.

4.1.1 Bases for the framework

The components of an SDI were outlined inter alia by (Strain et al., 2006) and (Rajabifard. & Williamson, 2001) (see subsection 2.1.3 on page 18). These publications showed that the two components people and data can be linked through the components standards, policies and access networks (see figure 4.1).

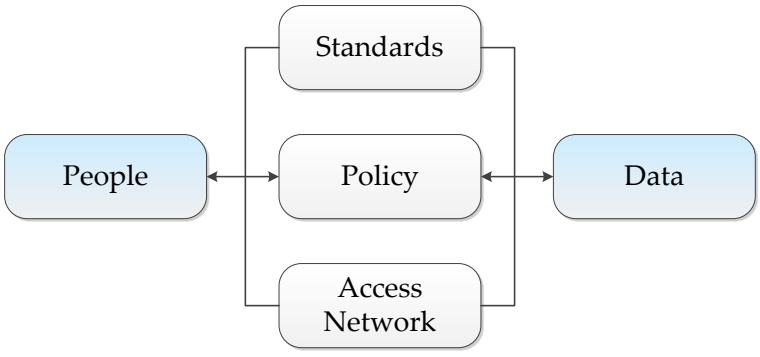


Figure 4.1: *Components of an SDI (modified after (Rajabifard. & Williamson, 2001))*

(Steudler et al., 2008) take these components and define areas that should be inspected when assessing spatial data infrastructures (SDIs): Policy Level (Policy), Management Level (Standards and Access Network), Operational Level (Access Network and Data), Other Influencing Factors (People) and Performance Assessment. For the defined areas, the paper then suggests possible indicators for the evaluation of SDIs (see table 4.1). The paper also outlines that *"[. . .] evaluation is about finding answers to questions such as 'are we doing the right thing' and 'are we doing things right' "*.

Table 4.1: *Possible indicators for Evaluating SDIs (Steudler et al., 2008)*

Level	**Area**	**Possible Indicators**
Policy Level	Policy	• existence of a government policy for SDI • handling of intellectual property rights, privacy issues, pricing • objectives for acquisition and use of spatial data
Management Level	Standards	• standardisation arrangements for data dissemination and access network • institutional arrangements of agencies involved in providing spatial data • organisational arrangements for coordination of spatial data • definition of core datasets • data modeling • interoperability
	Access Network	• access pricing • delivery mechanism and procedure • access privileges • value-adding arrangements
Operational Level	Access Network	• type of network • data volume • response time
	Data	• data format • data capture method • definition of core datasets • data maintenance • data quality and accuracy
Other Influencing Factors	People	• number of organisations and people involved • opportunities for training • market situation for data providers, data integrators, and end-users
Performance Assessment		• degree of satisfying the objectives and strategies • user satisfaction • diffusion and use of spatial data and information • turnover and reliability

(Najar et al., 2007) undertake a similar approach for the assessment of SDIs by proposing three components (data and metadata, web services, standards) accompanied by several indicators for each of the components. Since indicators could have either a technical or organizational meaning they are further classified by these two factors (see table 4.2).

Table 4.2: *Indicators for comparing SDIs on the basis of Web services and data management (Najar et al., 2007)*

Component	**Indicator**	
	Technical	*Organizational*
Data and metadata	1. Data capture process 2. Definition of core datasets 3. Data format and conceptual model 4. Data management 5. Data quality and accuracy 6. Common modelling language and tools 7. Harmonization of data and metadata	8. Custodianship 9. Data sharing and partnerships agreements 10. Business models 11. Coordinating arrangements
Web services	12. Application profile 13. Clearinghouse/Geoportal	14. Clearinghouse organization
Standards	15. Interoperability	16. Organizational arrangements for standardization

Because the aim is to evaluate and compare MSDIs instead of "regular" SDIs we may have to adjust the components outlined at the beginning (see figure 4.1). When examing (Butler et al., 2011) we discover that *"the four main components of a successful coastal and ocean information network (COIN), as an important component of an SDI, are:*

- *online access to data using recognized standards*
- *metadata catalogues that can be used to search for geospatial information*
- *a web interface that allows users to search, access and retrieve the best available information from the most reliable sources*
- *active participation of data providers and data users to ensure that the right data are available to contribute to more effective decision-making."*

These components are more or less equal to the five components a SDI consist of which were mentioned previously (see figure 4.1) in a slightly different form. When we compare them we can map them to:

- online access to data using recognized standards
  → people, data, standards and access networks
- metadata catalogues that can be used to search for geospatial information
  → people, data, standards, policies and access networks
- a web interface that allows users to search, access and retrieve the best available information from the most reliable sources
  → people, data, standards, policies and access networks
- active participation of data providers and data users to ensure that the right data are available to contribute to more effective decision-making
  → people, data, policies

With the preceding comparison it can be seen that the components of SDIs and MSDIs are very similar and comparable. Said comparability enables us to use the indicators (which were meant for SDIs) described in the two beforehand mentioned papers for the evaluation of MSDIs. At this point a remaining open question is: are any additional indicators required to assess MSDIs?

In order to see if any additional indicators are required for MSDIs we are looking at the definition of the term MSDI by (Russell, 2009). MSDIs are *"the component of a National SDI that encompasses marine and coastal geographic and business information in its widest sense. A MSDI would typically include information on seabed bathymetry (elevation), geology, infrastructure (e.g. wrecks, offshore installations, pipelines, cables); administrative and legal boundaries, areas of conservation and marine habitats and oceanography."* This means that core data sets still fit as an indicator but that the core data sets are obviously different for a MSDI. (Welle-Donker, 2010) adds to this definition by stating that MSDIs are special

SDIs because inter alia *"[. . .] seas and oceans do not stop at national boundaries [. . .]."* Seas and oceans do not only ignore national boundaries but the boundaries itself respectively their definition could be problematic because they are subject to change over time. This is why an indicator is proposed which checks whether the *definitions for shorelines and/or maritime zones* differ in varying MSDIs. Furthermore (Welle-Donker, 2010) states that *"[. . .] the collection of marine data is highly fragmented"* This is because there is a manifold of stakeholders in the marine domain. The different stakeholders might use different metadata standards which can be problematic when developing an SDI. This is why *metadata coordination* might be important.

Apart from these additional marine-specific indicators there still might be indicators missing which might be useful to evaluate MSDIs (as well as SDIs in general since these will not be marine-specific). In (Steudler et al., 2008) metadata cannot be found at all in the list of possible indicators although metadata is mentioned in the paper several times. (Najar et al., 2007) on the other hand list "Data and metadata" as a component so that some of the indicators listed under this component also apply to metadata and the indicator "harmonization of data and metadata" mentions metadata explicitly. But what is not mentioned in both papers is the *availability of metadata* in general and a *metadata catalogue* in particular (this enables inter alia automatic harvesting). Metadata of course is linked to real data which is usually downloadable. To evaluate if the data might be helpful, ahead of download, view services (Web Map Services [WMS]) are needed. When a user found out that a data set is useful it would be even easier for him (at least in some cases) if he was able to integrate the data on the fly (through a Web Feature Service [WFS]) without having to download it. For both use cases an indicator is needed that looks at the *availability of services*. Another aspect also not found in the existing approaches is the *architecture* of a (marine) SDI. This is interesting in order to better understand how other (marine) SDIs got to their infrastructures and how they are built.

4.1.2 Compiling the framework

When we compile the approaches described in subsection 4.1.1 we end up with several indicators (see table 4.3) which will be explained in detail in subsection 4.1.3. For clarity purposes all indicators are classified into the factors technical and organizational.

Table 4.3: *Possible indicators for the evaluation of marine spatial data infrastructures*[a]

Topic	**Indicator**	
	Technical	*Organizational*
A - Data	• ***Core datasets***	• *Degree of involvement of different agencies/institutions* • *Data modelling* (**Harmonization of data and metadata**)
B - Metadata	• Availability of metadata/ Metadata catalogue • ***Data quality and accuracy***	• Coordination
C - Services	• Availability of services • **Performance** • **Geoportal**	• *Access privileges*/**Custodianship** • *Value-adding arrangements*
D - Standards	• ***Interoperability***	
E - Modelling		• *Existence of a government policy for SDI* • Architecture • Definition shoreline/maritime zones • **Business models**

[a]sources are *italics* for *(Steudler et al., 2008)* and **bold** for **(Najar et al., 2007)**

4.1.3 Description of the indicators

The first area (**A**) covers the organizational and technical indicators regarding data. Because data is most important in a MSDI this is the first area looked at. The indicator *core datasets* describes the basic reference spatial data covered by a country's MSDI. The datasets which could be covered are as follows (refer to subsection 2.1.5 on page 20):

- Bathymetry
- Shoreline and other maritime zones like EEZ
- Marine Cadastre
- Coastal imagery
- Marine navigation
- Tidal benchmarks
- Benthic/Nature conservation habitats

The indicator *degree of involvement of different agencies/institutions* looks at the degree of involvement of different institutions by reason that a MSDI has to incorporate various datasets coming from a wide range of agencies/institutions and these will be listed for this indicator and – if possible – compared to the ones that were left out. *Harmonization of data and metadata* respectively *data modelling* are important (and expensive) steps to provide users with data of quality that they are able to understand, interpret and discover. Another factor to be considered here are data updates with which it has to be ensured that the same metadata fields are used for equal or similar data sets (e.g. when handling time series data sets which are very common in the marine domain).

For area **B** it is important that data is augmented by metadata to be able to discover it and to know what the data is about later on. The indicator *availability of metadata/a metadata catalogue (CSW)* looks at the availability of metadata and tries to answer the questions "is it searchable?", "how is it held?" and "is it available through a standardized catalogue interface?". Because we are in the marine field much data will be sensor data thus describing data quality and accuracy of the data is a big issue. The indicator *data quality and accuracy* tells if metadata is available that describes details of the measurements and their accuracy (if OGC's Observations and Measurements [O&M] standard is used for the metadata this field is already covered). In general, it would be wise to use internationally

approved standards and if needed build profiles to meet special requirements. For this purpose, indicator *coordination* is designed because metadata should be homogeneous inside a MSDI or at least appear homogeneous to users. This is needed for example when chemical or physical measurements are published because they need the same measurement units to be comparable directly (on a map for example). To achieve this the metadata fields do not have to be the same but a mapping is needed which requires previously mentioned coordination. There should be a central coordination unit dealing with implementing metadata rules (what standard to use, how to build a profile and so on).

Having data and metadata for the data is a good thing so far but having services to use data on the fly without having to send files back and forth would increase efficiency considerably. In addition offering services for data and metadata is an INSPIRE requirement. For that purpose area **C** (services and interfaces) deals with the availability of services, their performance and the presence of a geoportal. From an organizational standpoint access privileges and value-adding arrangements are important, too.

For interoperability it is important that you are able to get marine-related data into any application of your choice and to be independent of a geoportal alone. Thus the indicator *availability of services* lists all the available services categorized in Discovery, View, Download and Transformation services. Not to forget here is the availability of a gazetteer. When you want to work with the data provided by services it is important that the services meet certain criteria regarding response time or *performance* in general. Furthermore the system has to be able to cope with large data sets and there should be an update cycle which is well documented. The MSDI should have a central entry point to access its data which is a *geoportal*. It is important that there is a search functionality and a map viewer. A central portal as single-entry point is the best way for the users because they do not have to know of and visit other portals. In some way connected discovery services are good, too. Additionally the indicator *access privileges/custodianship* asks if there was a focus on a role model which deals with actors or stakeholders of the system when the MSDI was modelled. Another question is whether there are *arrangements* with the private sector (companies etc.) which *add value* to the infrastructure. For example the sea cable a telecommunications company has laid might affect certain species/populations and to study that you need the data from the company. So it would be great if you would have that data already available through the infrastructure.

All the areas so far affect standards in some way. How and to what extent answers area **D** which asks in its only indicator *interoperability* what standards are used and does their usage lead to better interoperability. However, it has to be stated that this is really hard to measure respectively is almost immeasurable. Because it cannot be said what a good standard is or that it is better to use many standards and worse if only a few are used. Another problem is the existence of different versions of standards because it cannot be said if the latest versions are better than older versions. Apart from that this indicator asks if the stakeholders of the infrastructures are involved in standardization processes or organizations.

The last area (E) focuses on various aspects important for MSDIs from an organizational viewpoint. It considers the *existence of a government policy for a MSDI* and thus answers the question if the government backs up the developments. However, since the government most probably funds the development of a MSDI the backing of the government should be implicit. To better understand how other marine initiatives got to their infrastructure and how they are built the *architecture* and in particular the underlying *business models* are examined, too. Area E also gazes at the varying *definitions for shorelines and/or maritime zones* in diverse MSDIs from a legal point of view which is why this is classified as an organizational indicator.

4.1.4 Assessment of the so far found indicators

After identifying all indicators which sound interesting for MSDI evaluation an approach is needed to verify that the found indicators really add value to the evaluation process, that information

about indicators is obtainable and so on. For this purpose SMART criteria were used to assess the indicators.

Introduction to SMART criteria

According to (Wood, 2011) SMART criteria are usually about the assessment of goals and targets. SMART is a mnemonic which most often stands for **S**pecific, **M**easurable, **A**ttainable, **R**elevant and **T**imely. The specific meaning of these terms can only be defined by knowing the field they get applied to and defining them for that particular topic. Since SMART can be applied to numerous areas, there are other meanings for some of the letters, inter alia[1]:

- A: achievable, accessible, added value, appropriate, actionable
- R: result-oriented, realistic, reasonable, reliable
- T: tangible

(Buglione & Abran, 2003) also maps the SMART criteria to the "5W's+H" rule (What, Why, Who, Where, When, How) making them easier to understand (see table 4.4).

Table 4.4: *Mapping between SMART criteria and the "5W's+H" rule (Buglione & Abran, 2003, p. 295)*

SMART criteria	5 W's+H element
S – Specific	What
M – Measurable	How
A – Add Value	Actionable & Why
R – Realistic	Who, Where
T – Timely	When

SMART criteria got applied in the marine area, too, mainly in the field of marine protected areas, marine conservation and planning objectives ((Lumb et al., 2004) and (Rice et al., 2005) for example) and marine protection targets (Wood, 2011). SMART criteria were also used for evaluation frameworks, e.g. for injury surveillance systems (Mitchell et al., 2009) or in software process improvement frameworks (Buglione & Abran, 2003). However, to the author's knowledge to date, SMART criteria were not applied to indicators for the evaluation of (marine) spatial data infrastructures. That is why the following elaborations are assessing the indicators which try to evaluate MSDIs.

(Buglione & Abran, 2003) is also pointing out that the SMART criteria not necessarily have to lead to the usage of Boolean values (true and false) but that also an ordinal scale could be used to assess indicators with SMART criteria (see figure 4.2) leading to an indicator assessment grid (IAG). If an ordinal scale is used for the assessment of the indicators an IAG has to be developed after defining the SMART criteria.

SMART	Description	0 *Poor/Absent*	1 *Fair*	2 *Good*	3 *Excellent*
S - Specific	*Indicators must be specific and targeted to the area intended to be measured.*	Not focused on the area intended to be measured	Informally addresses and covers the area intended to be measured.	Addresses and covers the area intended to be measured.	Properly addresses and covers the area intended to be measured.
M – Measurable	*Indicators must permit collection of accurate and complete data*	Incomplete or bad definition	Formal definition of the elements needed for	Definition of the elements needed for cal-	Complete and exhaustive definition of the elements for

Figure 4.2: *The Indicator Assessment Grid (excerpt (Buglione & Abran, 2003, p. 296)*

[1] cf. (Robinson et al., 2009), (Lumb et al., 2004), (Rice et al., 2005), (Buglione & Abran, 2003) and (Mitchell et al., 2009)

Defining the SMART criteria for the indicators

To verify that the selected indicators for the evaluation are indeed useful we have to set up criteria to assess them (see table 4.5). The "5 W's and H" rule is used to define the SMART criteria. Firstly we ask if *what* is measured is specific and unambiguous. After that we have to think about *how* or if the indicators are measurable. Then we have to think about *why* we include a certain indicator - we question if the indicator creates added value for the evaluation framework. Because most of the time the information about the MSDIs only stems from bibliographical sources we have to ask *who* gave information on a certain indicator and *where* this information was found. In the end we are asking about the reliability of indicators and the probability that reliable information will be found for the indicator. The question *when* is irrelevant when it comes to indicators for evaluation so time will be left out.

Table 4.5: *Selected SMART criteria and their meaning*

Criterion	**Description**
Specific	Indicators have to be specific and unambiguous.
Measurable	Indicators have to be measurable which means that grades can be given during the evaluation.
Add Value	Indicators have to create added value for the evaluation framework so that the MSDIs can be compared and evaluated better.
Reliable	Indicators have to be reliable meaning that information on a certain indicator is probably not restricted.

Table 4.6: *Indicator Assessment Grid*

Criterion	**Description**	**Indicator assessment**			
		0	*1*	*2*	*3*
Specific	Indicators have to be specific and unambiguous.	Indicator is ambiguous.	Indicator is rather specific, could lead to confusion.	Indicator is specific and unambiguous.	Indicator is totally specific and unambiguous.
Measurable	Indicators have to be measurable meaning that grades can be applied, thus them being rateable.	Indicator cannot be measured.	Indicator is very hard to grade.	Indicator can be measured.	Indicator can very easily be measured.
Added value	Indicators have to add value for the evaluation framework to improve comparison and evaluation.	Indicator does not create added value at all.	Indicator adds very little value.	Indicator creates added value.	Indicator adds much value.
Reliable	Indicators have to be reliable meaning that information on a certain indicator probably is not restricted.	The probability that reliable information could be found is nearly zero.	The probability that reliable information could be found is very low.	It is likely that reliable information could be found.	It is very likely that reliable information could be found.

Development of the indicator assessment grid
To assess the indicators an ordinal scale is used leading to an indicator assessment grid (IAG, see table 4.6). The IAG is used because a simple yes/no assessment for the indicators does not seem appropriate because of their complexity.

Applying the SMART criteria to indicators for the evaluation of MSDIs
The IAG is now used to assess the indicators (see table 4.7) by assigning values to each criterion which of course are very subjective estimates by the authors. The specific values will be set aside apart from one example – "availability of services" – because of their subjective nature and limited space to describe why each indicator was applied. This indicator is specific (*S*) as it just lists the offered services (grade 3), but it is very hard to measure (*M*) because there are so many services types and it cannot be defined how many have to be offered to get a ++, or a + and so on (grade 1). However, it does add value (*A*) to the evaluation because services are very important for users nowadays (grade 2) and because of that the available services are mostly well communicated to the users which means that this indicator should be reliable (*R*, grade 2).

Table 4.7: *Assessment of the indicators*

Criterion / **Indicator**	*S*	*M*	*A*	*R*	∅
Core datasets	2	1	3	3	2,25
Degree of involvement of agencies	3	1	3	3	2,5
Data modelling (Harmonization of data and metadata)	1	0	1	1	**0,75**
Availability of metadata/catalogue	3	1	3	2	2,25
Data quality and accuracy	2	1	2	1	**1,5**
Coordination	2	1	2	2	1,75
Availability of services	3	1	2	2	2
Performance	3	2	2	2	2,25
Geoportal	3	2	2	2	2,25
Access privileges/custodianship	2	1	2	1	**1,5**
Value-adding arrangements	1	1	1	1	**1**
Interoperability	2	1	2	2	1,75
Existence of a government policy for SDI	2	1	2	3	2
Architecture	2	0	2	2	**1,5**
Definition shoreline/maritime zones	2	0	1	1	**1**
Business models	2	0	1	1	**1**
Average	$\sim$ 2,19	$\sim$ 0,88	$\sim$ 1,94	$\sim$ 1,81	$\sim$ 1,70

After applying the IAG a threshold has to be defined so that it can be decided which indicators are kept and which are not used for the evaluation framework because they fail to add value and/or be specific and/or be measureable and/or be reliable. To compute the threshold T firstly a mean value ∅ of the four criteria was calculated for every indicator. After that the sum of all the mean values was calculated and then divided by the total number of indicators giving the total mean value of all the indicators (see formula (4.1)).

$$T = \frac{\sum_{i=1}^{N} \frac{S_i + M_i + A_i + R_i}{C}}{N} \tag{4.1}$$

where

N – total number of indicators	S_i – value of criterion *specific* for i
i_1 – First indicator (A1)	M_i – value of criterion *measurable* for i
i_N – last indicator (E4)	A_i – value of criterion *added value* for i
C – total number of criteria	R_i – value of criterion *reliable* for i

The result T = 1,7 (rounded) is the threshold used for the decision which indicators will be left out. Every indicator which has a higher average mean than 1,7 is kept and every one below will be scrapped. But there is one exception from that rule – the indicator *architecture*. Although this indicator is measureable by no means the architecture of a MSDI is still very interesting and will be kept just for the sake of including this kind of information in the evaluation.

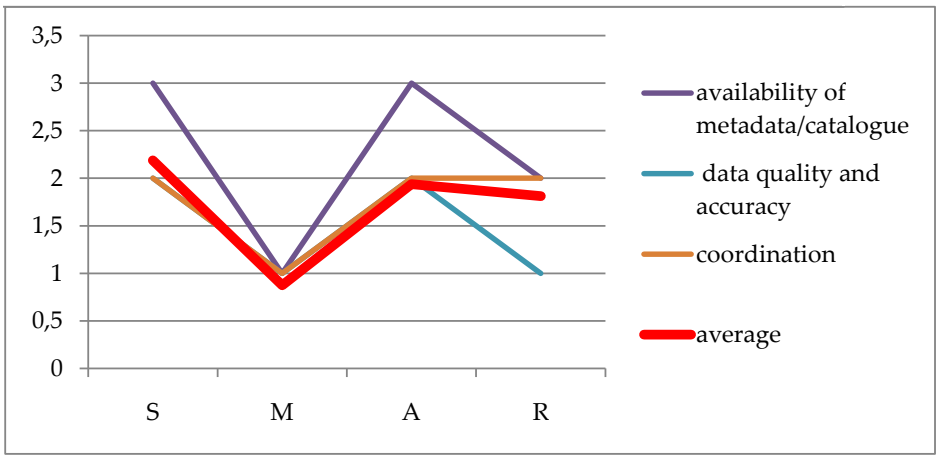


Figure 4.3: *Line chart of indicator assessments for area B*

This approach might seem complex at a first glance but when visualized with a line chart (see figure 4.3) underachievers can be easily seen because they clearly are below the average line which indicates the mean value of each criterion. When inspecting indicator *coordination* for example the line chart shows that it aligns closely with the mean value line but is slightly above it alltogether. This reflects its assessment mean value of 1,75 which is just a little over T (1,7) which is why this indicator will be included in the final evaluation framework. All the indicators which have a mean value below 1,7 (T) will not be included in it except for the indicator *architecture* as stated above. The reason for this is that information about the architecture of a MSDI is very likely to be found (thus the reliability rating of 2). However, this is not true for the other indicators which will be completely left out. On top of that architecture has the highest mean value of the indicators (amongst two others) scoring right below T which means that most of the other indicators left out are less specific, add less value and/or do not add as much value for an evaluation. Because architecture is a special case and cannot be measured no score will be assigned to it when using the evaluation framework. Information about the architecture will just be stated as text and/or figures.

The remaining indicators after applying the assessment are presented in table 4.8 and are the ones which section 4.2 uses for the evaluation of international case studies.

Table 4.8: *Remaining indicators for the evaluation of marine spatial data infrastructures*

Area	**Indicator**	
	Technical	*Organizational*
A - Data	1 Core datasets	2 Degree of involvement of different agencies/institutions
B - Metadata	1 Availability of metadata/ metadata catalogue	2 Coordination
C - Services	1 Availability of services 2 Performance 3 Geoportal	
D - Standards	1 Interoperability	
E - Modelling		1 Existence of a government policy for SDI

4.2 International case studies

After building the framework this section uses it to analyse and evaluate the efforts of selected existing marine SDIs that were introduced in section 3.2 on page 49. The efforts of Canada (Marine Geospatial Data Infrastructure [MGDI], COINAtlantic and GeoPortal, see subsection 4.2.4), the USA (CMSP Registry, Marine Cadastre, Data.gov and Geoplatform, see subsection 4.2.5) and Australia (AODN IMOS Ocean Portal and Australian Marine Spatial Information System [AMSIS] , see subsection 4.2.3) were selected because they were developed before 2004 (Williamson et al., 2004) and thus should be quite advanced and information is likely to be found instead of work in progress efforts. Ireland (Marine Irish Digital Atlas [MIDA]) is building its infrastructure for an equally long time and is particularly interesting because it is a European country which means that it should be affected by the same legislation as Germany (see subsection 4.2.1). The same is true for the United Kingdom whose efforts (MAGIC/CAMRA and MEDIN) are discussed in subsection 4.2.2. Lastly the developments in Germany respectively the outcome of the MDI-DE project are assessed in subsection 4.2.6.

Because the process of applying all the indicators to existing MSDIs takes up a great amount of space this procedure will be shown using one example. For the reasons stated above Ireland was selected as an example showing how to apply the framework to a MSDI and evaluate it with the framework. For the other MSDIs only a summary is provided here. The detailed evaluations are found in appendix B from page 163 on.

Since the evaluations are solely based on literature and the portals of the MSDIs the results of the evaluations are subjective to a certain degree. The evaluation scale is:

++ very good, + good, $\pm$ not applicable, $-$ not so good, $--$ bad

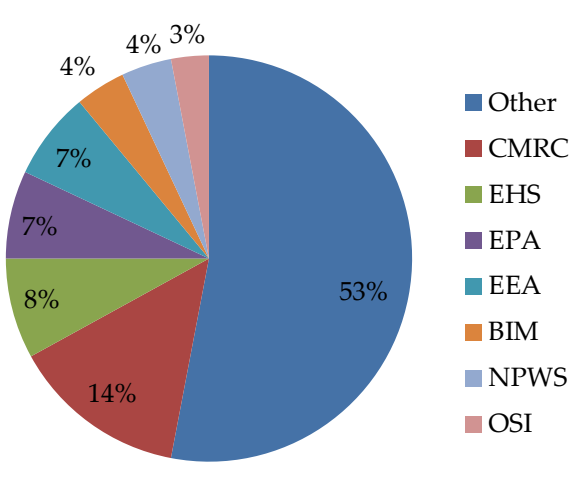
4.2.1 Ireland: Marine Irish Digital Atlas

Ireland's Marine Irish Digital Atlas (MIDA) which was introduced in subsection 3.2.3 will be assessed by using the evaluation framework looking at each indicator in a single paragraph. The evaluation results in table 4.9 which shows the findings at a glance and a short summary.

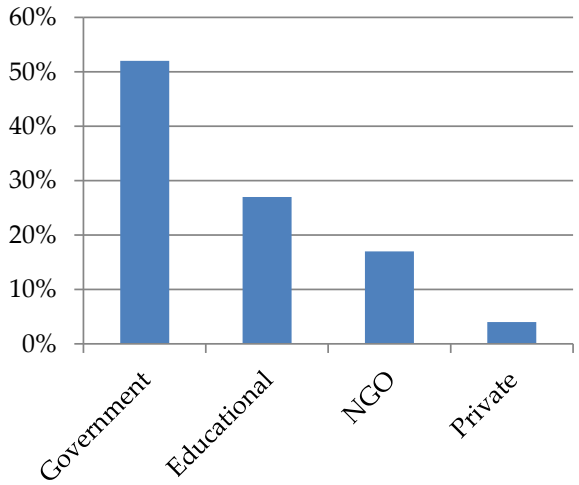
A – Data

A1 – Core datasets Initially MIDA faced two problems when trying to acquire (core) data sets. On the one hand it was challenging to get to know who has what data and on the other hand some data owners did not want to contribute to MIDA because of the commercial nature of their data or because they were developing their own GIS. But fortunately some key data owners (such as the Marine Institute and the Department of Communications, Energy and Natural Resources) were aware that data should not be locked up somewhere but rather be included in an infrastructure (O'Dea et al., 2009).

According to (O'Dea et al., 2007) we find these – not limited to core – data sets: physical environment, coastal habitats, imagery, biology, management, human impact, conservation, environmental monitoring, infrastructure, industry, culture & heritage, natural resources, fisheries, aquaculture & agriculture and tourism & recreation.



(a) The primary data supplier organizations (n=7) according to percentage of data layers provided to the Atlas. All those that contributed less than 3 % of datasets are grouped as other (n=28) (modified after (Dwyer et al., 2011))



(b) Sectors to which data supply organizations (n=35) belong (modified after (Dwyer et al., 2011))

Figure 4.4: *Data suppliers for MIDA*

A2 – Degree of involvement of different agencies/institutions Figure 4.4 shows that there are 35 organizations providing data for MIDA of which more than half are government organizations like the Environmental Protection Agency (EPA, 7 %), the Coastal & Marine Resources Centre (CMRC, 14 %) and the Bord Iascaigh Mhara (BIM, Irish Sea Fisheries Board, 4 %). 27 % of the data sets are supplied by educational institutions while 17 % come from non-governmental organizations (NGO's) and only 4 % are from the private sector. With such a variety of data sources all fields of marine and coastal data should be covered in Ireland very well (Dwyer et al., 2011).

Connection to EU directives Because Ireland as a European country is affected by EU directives just like Germany it is checked how Ireland complies with these directives. This does not qualify as an indicator because it is not general enough which means that it can be applied only to European infrastructures to include it in the evaluation framework. However, this information is important as pointed out and fits best into the **data** topic. Some organisations which contribute to MIDA have to publicise data for the EU's Water Framework Directive. One of them is the Environmental Protection Agency (EPA) which contributes a few layers to MIDA: lakes and rivers, bathing water quality at

beaches, coastal and transitional waters and river basin districts (O'Dea et al., 2009). Since ISO 19115 is the core of the metadata standard for INSPIRE MIDA to some extent ties in the INSPIRE initiative, too.

B – Metadata

B1 – Availability of Metadata/Metadata catalogue (CSW) MIDA does provide metadata for all its data sets in ISO 19115 form (see indicator B3 on page 91 for further information) but does not offer a service for discovering these. But outside of MIDA there is Marine Data Online (MDO) which is a GeoNetwork-based data catalogue developed and hosted by the Marine Institute offering an INSPIRE compliant data discovery service (http://catalog.marine.ie). Although the Marine Institute has a large amount of data it does not cover all marine data in Ireland. Because of that the Irish Spatial Data Exchange (ISDE, http://www.isde.ie) was started in 2004 involving other agencies as well. ISDE offers a CSW 2.0.2 and is also able to search the catalog of the Marine Institute.

B2 – Coordination The Irish Organisation for Geographic Information (IRLOGI) is the umbrella organisation for the geographical information sector in Ireland. Its Geo-ID (Geospatial Information Directory) initiative has evaluated metadata standards of which ISO 19115 was found to suit the needs of Irish spatial data best. IRLOGI then developed a metadatabase through which organisations could publish their metadata, however, update and maintenance of this portal has been problematic and unsteady from the start. In 2002 the Irish Government initiated the Irish Spatial Data Infrastructure (ISDI) and the Department of Environment, Heritage and Local Government began developing it. This department – among other things – developed an ISO 19115 metadata profile which all parties in the ISDI community (including the marine community with organisations like EPA, the Marine Institute etc.) adopted.

That is why ISO 19115 was chosen for MIDA. Because of the multitude of organisations contributing data to MIDA ISO 19115 was reduced to a profile with 55 elements (a complete listing can be found in appendix B of (Dwyer et al., 2011)) including all obligatory elements, some optional ones concluded to be important for marine data and some not included in the ISO standard (bounding box in local projection for example). For user information three levels of metadata were chosen:

- *Abstract Metadata*: simple description or explanation of each dataset, first level a user sees when viewing metadata with a link to the Discovery Metadata, selection of elements stored in the MIDA XML database
- *Discovery Metadata*: contains the main metadata fields, all datasets displayed in the MIDA have this Discovery Metadata, if more information is available there is a link to Full Metadata, stored in the MIDA XML database
- *Full Metadat*a: supplied by the data owner, availability depends on whether or not it exists, and its quality is the responsibility of the data owner, not stored in the MIDA XML database

For a user-friendly presentation XSLT stylesheets are used to display the queried metadata records which the XML database supplies. Before display of metadata of course comes entering metadata into MIDA. For that purpose several tools were evaluated like the ESRI ArcCatalog ISO wizard, M^3Cat, Enraemed and the ANZLIC (Australia New Zealand Land Information Council) Metadata Collector. But due to too many offered functions and thus complexity of the tools a customized data entry wizard was developed using HTML, JavaScript and Java Server Pages (JSP). After the user submits metadata through HTML forms a XML document will be saved on the server which the administrator later collects, reviews and inserts into the database (O'Dea et al., 2004b) and (Dwyer et al., 2011).

C – Services and Interfaces

C1 – Availability of Services For the provision of web services for MIDA several tools were evaluated including ESRI ArcIMS, Intergraph GeoMedia WebMap and MapServer. For the evaluation the factors range of functionality, ease of customisation, dependencies on other software and cost were taken into account. Based on these MapServer was chosen because a complex GIS was not necessary and the core functionality of MapServer was sufficient for map visualisation. Besides that MapServer is relatively easy to customise using scripting languages like JavaScript, Perl, PHP and Java and MapServer is free open source software which means it is a relative low-cost and long-term solution. Due to the fact that MapServer is used MIDA supports (and offers) data with OGC services like WMS, WFS and WCS.

Shown in figure 4.5 is MIDA's ability to present time series data originating from services (Dwyer et al., 2003), (Dwyer et al., 2011) and (O'Dea et al., 2004a).

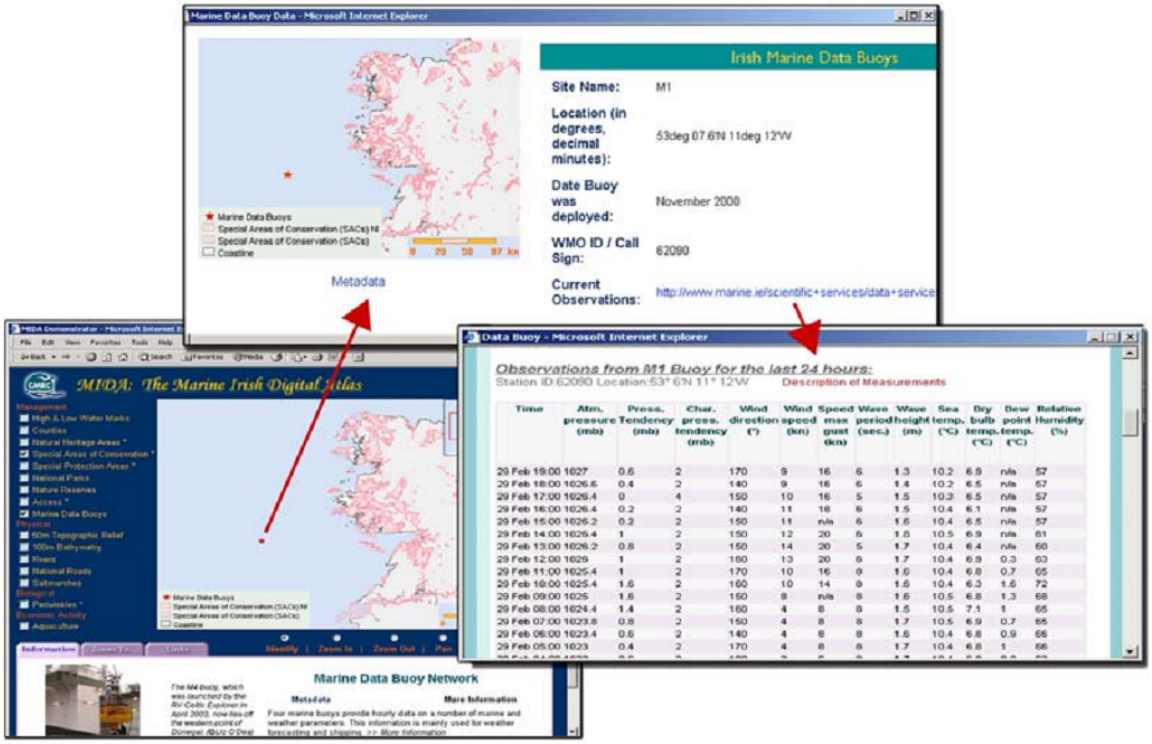


Figure 4.5: *When available, the MIDA links to near real time data, such as the observations by marine data buoys (O'Dea et al., 2004a)*

C2 – Performance Unfortunately no information could be found regarding performance.

C3 – Geoportal MIDA's portal (or "atlas" as they call it, available at mida.ucc.ie) is a fully functional web GIS offering common functionalities such as search, add/remove layers, zooming and so forth. What is more special to the atlas is the introductory information displayed below the map when a layer is selected (see figure 4.6). Besides the introductory information found there the user is also able to access the layer's metadata and get more information.

D – Standards / D1 – Interoperability Regarding metadata the usage of an ISO 19115 profile has already been discussed in detail in indicator B3 on page 91 and the support of OGC web services standards has been pointed out in indicator C1 on page 92. That is why accessing data and thus interoperability is not an issue in Ireland's marine community anymore.

Figure 4.6: *Example of introductory information found in MIDA's atlas (mida.ucc.ie)*

E – Modelling

E1 – Existence of a government policy for SDI The topic involvement of the government was touched in indicator B2 on page 91 where it was pointed out that the ISDI initiated its development. (Bartlett et al., 2004) pointed out that there was the *"[. . .] recognition that ISDI should seek to fully incorporate marine and coastal data from the very beginning [. . .] "*.

E2 – Architecture As already pointed out MIDA is using an XML database for metadata and MapServer for map delivery and offering services. (Dwyer et al., 2011) states that *"[. . .] a geospatial database was considered unnecessary [. . .] "* thus geospatial data is stored in the file system (Shapefiles for vector data and GeoTIFF for imagery/raster data, shown as *geospatial data repository* in figure 4.7). Data preparation tasks like conversion, cleaning, reprojection and attribute table editing were mostly done in ESRI ArcGIS. But for tasks increasing MapServers performance like tiling large raster layers and optimizing vector display other tools were used. For the former task gdaltindex and for the latter shptree were used.

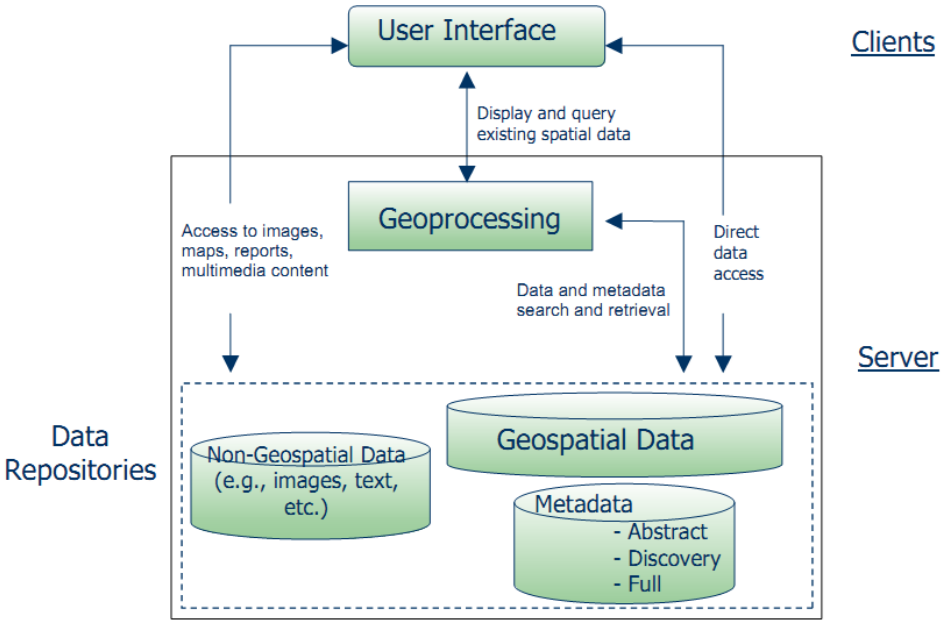


Figure 4.7: *Elements that make up the MIDA system based on the INSPIRE model, 2002 (Dwyer et al., 2003)*

Summary In conclusion (see table 4.9) the Irish began early and did very well on implementing a marine SDI. Through its embedment in the Irish Spatial Data Infrastructure (ISDI) with incorporation of marine and coastal data from the very beginning the government supported MIDA. Furthermore governmental agencies like the Marine Institute and the Department of Communications, Energy

and Natural Resources provided many core data sets. Through the use of the ISO 19115 standard – respectively the metadata profile developed by the Department of Environment, Heritage and Local Government – aspects of data quality and accuracy got implemented. Through the development by the Department of Environment, Heritage and Local Government and the initiation of the ISDI metadata was coordinated from a central point from a government level. MIDA emphasizes the use of free open source software and offers its services through MapServer and – unfortunately through another portal – its metadata through GeoNetwork. On the negative side it has to be stated that although a fully-functional portal is available for MIDA the look and feel of it could be improved (e.g. there is always a new window for adding/removing layers).

Table 4.9: *Results of the evaluation of Irish efforts*

Area	**Indicator**	
	Technical	*Organizational*
A - Data	1. ++ (Core datasets)	2. ++ (involvement different agencies) 3. ± (Connection to EU directives)
B - Metadata	1. + (Availability of metadata/catalogue)	2. ++ (Coordination)
C - Services	1. ++ (Availability of services) 2. ± (Performance) 3. + (Geoportal)	
D - Standards	1. ++ (Interoperability)	
E - Modelling		1. ++ (Existence of government policy)

4.2.2 United Kingdom: MAGIC/CAMRA and MEDIN

The UK efforts that were introduced in subsection 3.2.4 on page 52 perform *very good* in indicator A1 – as depicted in table 4.10 – mainly because MEDIN is offering a wide range of core data sets. Because many organizations are involved in both initiatives the UK performs *very good* in A2, too. As will be seen in D1 MEDIN is compliant with INSPIRE and there are also plans to report through MEDIN for the WFD which means the UK is keeping an eye closely on EU initiatives and results in *very good* in indicator A3. While much metadata is available for both initiatives there is no CSW for either one of them resulting in a *good* rating for B1. Both portals do not offer (OGC) web services on their own but since it is unknown if the data holders which MEDIN is linking to offer services indicator C1 together with C2 (performance) cannot be evaluated. MAGIC/CAMRA is offering a fully-functional geoportal with all necessary tools which MEDIN is not. MEDIN just provides a map in its map search thus resulting in a *good* for C3. Standardswise there only is very outdated information available for CAMRA but it mentions ISO metadata standards which led to the assumption that they sought interoperability through the usage of standards. MEDIN metadata is compliant with ISO 19115 and INSPIRE and uses ISO 19139 as schema set for implementing ISO 19115 resulting in a *very good* in the field of interoperability (D1). Many governmental organisations are involved in both initiatives which means that the government is backing up these developments (*very good*) in E1.

Table 4.10: *Results of the evaluation of British efforts*

Area	**Indicator**	
	Technical	*Organizational*
A - Data	1. ++ (Core datasets)	2. ++ (involvement different agencies) 3. ++ (Connection to EU directives)
B - Metadata	1. + (Availability of metadata/catalogue)	2. ± (Coordination)
C - Services	1. ± (Availability of services) 2. ± (Performance) 3. + (Geoportal)	
D - Standards	1. ++ (Interoperability)	
E - Modelling		1. ++ (Existence of government policy)

4.2.3 Australia: AMSIS and Ocean Portal

The two Australian approaches – the Australian Marine Spatial Information System (AMSIS) and the Integrated Marine Observing System (IMOS) Ocean Portal – were introduced in subsection 3.2.1. According to the evaluation results shown in table 4.11 (the in-depth evaluation can be found in subsection B.1) it can be concluded that Australia put much effort into their approaches and by providing the framework – the Australian Spatial Data Infrastructure (ASDI) – managed to develop a broad MSDI. Just like the Canadian MSDI the Australian one is missing a common entry portal. Users have to look at least at two different portals (and of course have to be aware of their existence). On the positive side Australia strongly focuses on free open source software (GeoNetwork etc.) and free and open data usable by anyone without restriction. Furthermore they divided their system into several nodes making it better manageable and scalable.

Table 4.11: *Results of the evaluation of Australian efforts*

Area	**Indicator**	
	Technical	*Organizational*
A - Data	1. ++ (Core datasets)	2. ++ (involvement different agencies)
B - Metadata	1. ++ (Availability of metadata/catalogue)	2. ++ (Coordination)
C - Services	1. ++ (Availability of services) 2. + (Performance) 3. ++ (Geoportal)	
D - Standards	1. + (Interoperability)	
E - Modelling		1. + (Existence of government policy)

4.2.4 Canada: Marine Geospatial Data Infrastructure (MGDI), COINAtlantic and GeoPortal

There are even more portals in the marine domain in Canada than in Australia. There are at least three of these which were discussed in subsection 3.2.2 (in-depth evaluation in subsection B.2). The results of the evaluation of the Canadian efforts can be seen in table 4.12 indicating that Canada's approaches perform very well in area A because of its many core datasets and the broad variety of involved agencies/institutions. In area B a minor issue is that there is no central catalogue available which is desirable. But apart from that there is much metadata available in catalogues and everything is well organized with recognition of international trends in standardization. Area C is where the most points are lost because there could be more services available and a single central geoportal is lacking. This means that although there are portals for their initiatives users have to know of all the initiatives and portals and have to visit each of them to look for and get data. Unfortunately nothing really can be stated for C2 (performance). The rating in area D and E is overall very good due to the facts that the CGDI is endorsing and/or investigating a multitude of standards and that the CGDI is the national spatial data infrastructure (NSDI) of Canada which means that it is implemented by the Canadian government and that the CGDI " [. . .] recognizes that governments have a responsibility to make geospatial information available [. . .]." (Labonte et al., 1998). Thus the government backs up the Canadian MSDI developments. The only problematic area to be seen is the division into several projects and thus missing a central entry point for marine data so that users would not have to look at several places to get the data they need.

Table 4.12: *Results of the evaluation of Canadian efforts*

Area	**Indicator**	
	Technical	*Organizational*
A - Data	1. ++ (Core datasets)	2. ++ (involvement different agencies)
B - Metadata	1. + (Availability of metadata/catalogue)	2. ++ (Coordination)
C - Services	1. + (Availability of services) 2. ± (Performance) 3. + (Clearinghouse/geoportal)	
D - Standards	1. ++ (Interoperability)	
E - Modelling		1. ++ (Existence of government policy)

4.2.5 United States of America: CMSP Registry, Marine Cadastre, Data.gov and Geoplatform

The portals and platforms of the USA perform very well in indicator A1 (as can be seen in table 4.13) because most of the four approaches offer data inter alia about bathymetry, coastline and other zones such as EEZ (exclusive economic zone), marine cadastre, coastal imagery, marine navigation, tidal heights or nature conservation zones. The degree of involvement of different agencies/institutions (A2) is satisfactory, too. However, CMSP "just" offers its own data sets and the Geoplatform usually accesses Data.gov.

Metadata availability (B1) is very good for Data.Gov and CMSP. (Maali et al., 2010) showed that metadata is available for 95 % of the data sets in Data.Gov. However, since registration is required

this indicator could not be checked with Geoplatform and the Marine Cadastre even did not offer any metadata at all. Furthermore there is no (central) catalogue available. Since the FGDC is involved in almost all the approaches and developed metadata standards since the early 1990's and its influence on ISO standards (like ISO 19115) shows outstanding coordination in the metadata field.

There are many services available (C1) for all the efforts except Geoplatform (see the detailed evaluation in subsection B.3 on page 175 for an explanation) and the web application Service Status Checker by the FGDC (see 3.6.3.3 on page 72) shows that the US-American approaches are very aware of performance being important (C2). While Data.gov and Geoplatform are fully featured portals CMSP seems quite experimental with its ArcGIS.com map viewer and the Marine Cadastre lacks a search functionality.

With the many services, various download formats and the involvement of the FGDC for standardization the interoperability is very good. Because of the *National Ocean Policy Implementation Plan* and its National Ocean Council which proposed the focusing on Data.gov as central portal[2] a government policy is clearly existing.

Table 4.13: *Results of the evaluation of US-American efforts*

Area	**Indicator**	
	Technical	*Organizational*
A - Data	1. ++ (Core datasets)	2. + (involvement different agencies)
B - Metadata	1. + (Availability of metadata/catalogue)	2. ++ (Coordination)
C - Services	1. ++ (Availability of services) 2. ++ (Performance) 3. + (Geoportal)	
D - Standards	1. ++ (Interoperability)	
E - Modelling		1. ++ (Existence of government policy)

4.2.6 Germany: MDI-DE

The MDI-DE portal offers a broad variety of core data sets covering inter alia Bathymetry (BSH Nauthis, WMS Hydrography and WMS Topography) and Marine Cadastre (BSH WMS CONTIS Administration and Facilities). The MDI-DE project comprises eleven project partners and twelve associated partners covering every institution handling marine related data in Germany, three of them being federal agencies. MDI-DE focused on the Marine Strategy Framework Directive (MSFD) and set up feature and map services for it and also followed developments for the Water Framework Directive (WFD) and INSPIRE. These facts are reflected in the highest scores possible for every indicator in area A. Metadatawise MDI-DE is also very satisfactory because there is extensive metadata available for all data sets based on standards and also a central metadata catalog offering a CSW interface. That there are three working groups occupied with metadata harmonization completes the overall good impression of area B. With around 30 WMS and some WFS available (and a gazetteer at the LKN) which although being in non-productive environments offer satisfactory performance and conformity MDI-DE offers much data through services. However, MDI-DE is not the only entry

[2] http://www.whitehouse.gov/sites/default/files/microsites/ceq/national_ocean_policy_draft_implementation_plan_01-12-12.pdf

point for marine data in Germany which is the only negative point about its portal but nevertheless results in a "good" rating. It already became evident that many standards are used. These include ISO19115, ISO19119, INSPIRE for metadata and the OGC standards (CSW, WMS, WFS [Gazetteer]) and result in the best rating for area D. The fact that MDI-DE was funded by the BMBF (German Federal Ministery of Education and Research) and the national coordination (through the German Federal Ministry for Transport, Building and Urban Development [BMVBS]) show the existence of a government policy for a marine SDI and results in the best rating here, too.

Table 4.14: *Results of the evaluation of German efforts*

Area	**Indicator**	
	Technical	*Organizational*
A - Data	1. ++ (Core datasets)	2. ++ (involvement different agencies) 3. ++ (Connection to EU directives)
B - Metadata	1. ++ (Availability of metadata/CSW) 2. + (Data quality and accuracy)	3. ++ (Coordination)
C - Services	1. ++ (Availability of services) 2. + (Performance) 3. + (Clearinghouse/geoportal)	4. ++ (Custodianship)
D - Standards	1. ++ (Interoperability)	
E - Modelling		1. ++ (Existence of government policy)

4.3 Conclusions

This chapter firstly introduced existing evaluation frameworks for SDIs. These existing ones were used as a base to formulate an evaluation framework for MSDIs and were adapted to the special needs of a MSDI. The SDI evaluation indicators could be used because the components of a SDI and MSDI are comparable and similar. However, it was argued that MSDIs are still special SDIs which is why the indicators had to be adapted and later on had to be expanded with marine-specific indicators.

The resulting framework at this point was very subjective and thus a way was sought that could add a degree of objectiveness. This was found in SMART criteria which analysed each so far found indicator with the four different criteria *specificity*, *measurability*, *added value* and *reliability*.

After the framework was set up and assessed with SMART criteria it was applied to existing MSDIs which proved the usability, usefulness and feasibility of the evaluation framework. However, it has to be stated that evaluation was solely based on literature and the portals themselves which means that the results will have to be proved by marine experts (although so far paper reviews suggest that the results are fine).

5 Selected implementation aspects of an interoperable architecture

> *"'Interoperability' means the possibility for spatial data sets to be combined, and for services to interact, without repetitive manual intervention, in such a way that the result is coherent and the added value of the data sets and services is enhanced."*
>
> (INSPIRE, 2007, Art. 3 (7))

The MSDI of Germany (MDI-DE) was built mainly by commercial software development, in particular the portal. The infrastructure nodes on the other side were set up by project members on their own. Lessons learned from other MSDIs (section 5.1) contribute both to the requirements of the portal and the set-up of the infrastructure nodes. Especially for the infrastructure nodes requirements and rules are laid down in the reference model that is introduced in section 5.2. Before infrastructure nodes could be built it has to be clear what data sets and services already exist. In order to get an overview a database model had to be developed (section 5.3). On top of this model web forms and tables were implemented that incorporated deep knowledge of the marine domain (e.g. to be able to specify relations to directives) and the project itself. Now that the services are known and additional ones were set up at the infrastructure nodes their performance and availability become crucial. This is all the more important because the INSPIRE directive proposes requirements regarding these two features. This is why section 5.4 uses existing tools to monitor and evaluate services and tries to find out if the results of the tools are comparable and if the INSPIRE requirements can be evaluated with such tools or if its requirements are too ambiguous. Section 5.5 builds on the preceding section and implements a website that allows the visualization of the results of a specific monitoring tool. Words and thus vocabularies are an important aspect of SDIs, especially for marine SDIs because they are more scientifically oriented. However, pre-existing vocabularies in the marine domain in German are maintained poorly in regards to collaborative work and web availability. Thus there is the need to convert them, merge them and make them available through interfaces so that they can be used for other applications. Section 5.6 defines the requirements to achieve these goals. The actual implementation is laid down in section 5.7.

5.1 Lessons learned from (M)SDIs

Section 3.2 introduced some of the marine SDIs already existing and chapter 4 evaluated them based on the developed framework. The findings there in conjunction with the discoveries of section 3.3 (reference models for SDIs) are the base for this section. It checks which aspects are important and may be used for an own approach and to what extent. Subsection 5.1.1 analyses the existing reference models for SDIs and extracts useful characteristics. Subsection 5.1.2 analyzes the usage of UML in order to build and model an infrastructure while subsection 5.1.3 considers general architectural aspects found in other SDIs and reference models. Subsection 5.1.4 merges the findings of the three mentioned sections and expresses precise requirements for an own approach.

5.1.1 Use of RM-ODP

As already outlined in subsection 2.3.1 on page 29 RM-ODP offers a standardized framework to describe distributed applications in open and heterogeneous system environments by providing five

viewpoints (enterprise, information, computation, engineering and technology). Using RM-ODP simplifies the understanding of SDIs, i.e. how they are built and structured. This fact and and that RM-ODP is an international standard were reasons for other infrastructures like the Water Resources Observation Network (WRON, see subsection 3.3.2) and approaches to SDI design like *seamless SDI design* found in (Vaez & Rajabifard, 2012, p. 213) to use RM-ODP to define their structure.

However, many reference models of existing infrastructures that were examined in section 3.3 are not based on RM-ODP and are not using viewpoints. All of the German infrastructures introduced in subsection 3.3.1 use sub models instead of the very similar RM-ODP viewpoints. How their sub models can be linked to RM-ODP's viewpoints and that the sub models cover all viewpoints is shown in figure 5.1. Especially the reference model of GDI-NRW has a corresponding sub model for every RM-ODP viewpoint.

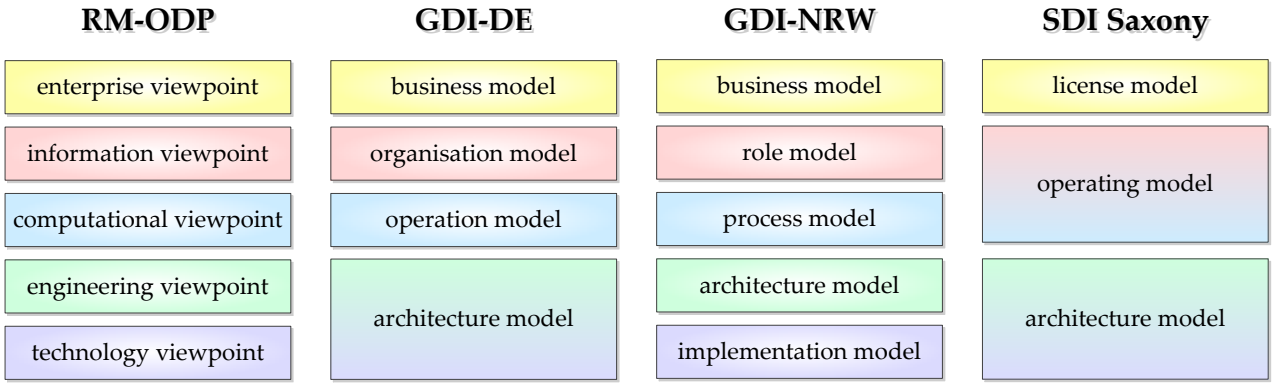


Figure 5.1: *RM-ODP's viewpoints and the corresponding sub models of selected German reference models*

5.1.2 Use of UML

Subsection 2.3.3 introduced UML modeling as an important technical aspect of a reference model. However, the reference models of the GDI-DE and SDI Saxony do not make use of UML. But in an outdated version of the reference model of the GDI-NRW (Kuhn et al., 2000) UML diagrams can be found as use case diagrams (e.g. to describe a user buying geodata, see figure 5.2), component diagrams (for the technical components of the SDI, see figure 3.5 on page 55), activity diagrams (to describe processes), interaction diagrams (to describe services) and sequence diagrams (e.g. client, server and data storage interaction).

Another existing reference model is using UML to model certain aspects of its infrastructure – the Water Resources Observation Network Reference Model (WRON-RM, see subsection 3.3.2). It uses use case diagrams inter alia to model the participating roles from an end users perspective or the discovery and access of dam data and class diagrams to model the data itself. All these diagram types can also be found in RM-ODP's viewpoints (namely its computational, engineering viewpoint and technology viewpoint) except for use case diagrams. However, these play an important role in the "4+1" View Model of Software Architecture (see subsection 2.3.2). Summing up it can be stated that these diagrams and UML in general play an important role when modeling an architecture because they are specified by standards and are used in real world examples. On top of that because UML is standardized it is easy to exchange models and diagrams.

5.1.3 Architectural aspects found in other infrastructures

An aspect regarding integration prominent in many countries is that many of the approaches found in a country tie in with a superordinate infrastructure (Australian approaches are subordinated to the Australian Spatial Data Infrastructure [ASDI], Canadian approaches to the Canadian Geospatial

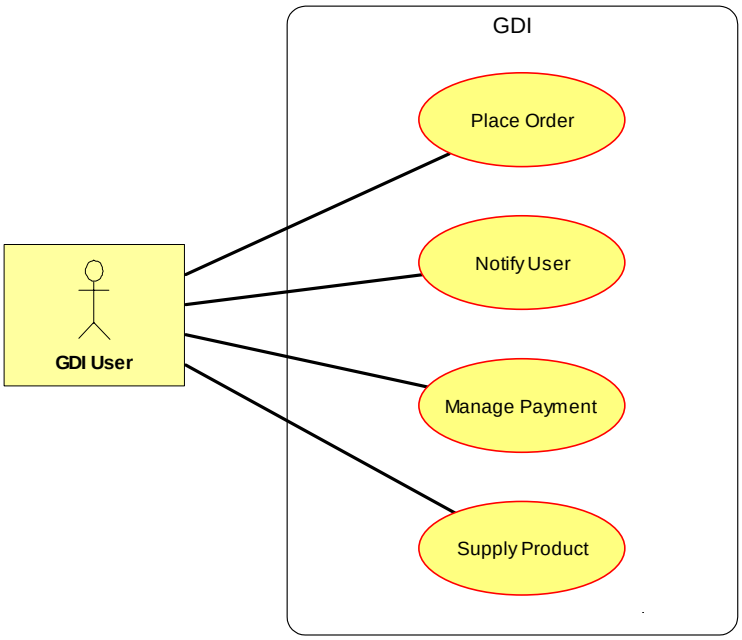


Figure 5.2: *GDI-NRW Use Case Diagram of a user buying a service (Kuhn et al., 2000, p. 12)*

Data Infrastructure [CGDI] and so on) except for Ireland where MIDA seems to be a stand-alone development.

Although not all of the approaches tie in with a superordinate infrastructure all the examined reference models are service oriented or at least promote the use of services. The services every reference model suggest are OGC's web map and feature services. The use of these OGC standards (as well as the ISO 191XX series and so on) emphasises that international standards play an important role for the construction of an SDI. Some of the reference models like DERM, GDI-NRW (suggested in (Bernard, 2002)) and the SDI Saxony suggest a catalogue service to publish metadata about the data that can be found in the SDI. DERM and the SDI Saxony also propose the use of layers to describe their architecture. Both use three layers. DERM suggests the three layers *clients*, *middleware* and *servers* while the SDI Saxony proposes the layers **application**, **service** and **data**. At a first glance these layers do not look similar. But because DERM's *clients* layers is about user **applications**, its *middleware* layer handles **services** and metadata and its *servers* only include various content repositories (**data**) all the layers can be mapped to each other and are in fact comparable.

5.1.4 Resulting requirements to construct an architecture

Existing SDI implementations and especially MSDI approaches offered valuable insight that will be considered for a novel approach.

RM-ODP turned out to be a useful framework to describe distributed applications in open and heterogeneous system environments. Because of this it suits the needs when building an SDI. The structure of an SDI becomes clearly visible through its viewpoints which offer in-depth insights into certain aspects of the architecture. RM-ODP's viewpoints can be directly mapped to the submodels of GDI-NRW as figure 5.1 shows. Because the submodel approach of GDI-NRW offers a better overview and separation of the aspects of the infrastructure it is favoured over the viewpoints of RM-ODP.

As subsection 2.3.3 on page 31 pointed out RM-ODP's viewpoints (and thus the submodels of GDI-NRW) can be constructed or enriched with the help of **UML** models and specific diagram types. UML was used in many reference models and is an international standard. This is helpful when building a complex system like an SDI because it is easy to exchange models and diagrams which is important when there are many participants bringing the reference model further.

Besides the use of RM-ODP and UML some specific **architectural aspects** can be adopted for a comptrehensive architectural blueprint. Almost all the approaches looked at, tie in with a superordinate infrastructure which means that the own approach has to tie in with the German Spatial Data Infrastructure (GDI-DE) which in turn ties in with INSPIRE. All the approaches are service oriented or at least use services based on standards (OGC and ISO). Some underline the importance of a catalogue service which is used to publish metadata about the data which can be found in the SDI. Thus the infrastructure to be built relies on services and should include a catalogue service. Some approaches use layers to describe the architecture (roughly there is an applications, services and data layer) which could also be considered for the own approach.

The last but nonetheless very important point found in (Altmaier et al., 2003, p. 5) is that a reference model has to dynamically adapt technical and market oriented developments – it has to be developed further continuously.

5.2 Reference model

In order to specify a structure for MDI-DE a reference model was built as a common effort of the project members[1] based on the standards introduced in section 2.3. The decision for RM-ODP was based on what was learned from other SDIs (see subsection 5.1.1). Furthermore the preceding section (section 5.1) influences and contributes to the development of the reference model, inter alia with architectural aspects learned.

One of these aspects (see subsection 5.1.1 and section 3.3) is that it makes sense to map the viewpoints of RM-ODP to specific submodels. Because a one to one matching between the viewpoints of RM-ODP and the submodels of the GDI-NRW is possible the reference model for MDI-DE was based upon GDI-NRW's reference model (see figure 3.4 on page 55).

5.2.1 Composition

The aims, requirements and interests of the participants towards the infrastructure are defined in the business model (see 5.2.1.1) which corresponds to the *enterprise viewpoint* of RM-ODP. Certain scenarios – which are part of the process model and are modeled as workflows – are the base for modeling the business processes. The participants (or actors) and their roles inside the system are collected in the role model (see 5.2.1.2) so that one can see if a certain actor holds what types of data for example. The flow of the scenarios and the state of the data inside the processes is described in the process model (see 5.2.1.3) which corresponds to the *engineering viewpoint* of RM-ODP. This model divides the process in activities and shows the course of actions and the state of data. Since the processes can change due to changing requirements the process model is dynamic as figure 3.4 on page 55 underlines. It is being modeled by activity and sequence diagrams. The architecture model (see 5.2.1.4) corresponding to the *informational* and *computational viewpoints* of RM-ODP characterizes the technical components (e.g. services, interfaces and clients) as well as their functions and the interrelations between them. In contrast to the process model the architecture model is static because it is defined at the beginning and the development orientates itself on this model because it is the base of the infrastructure. Since the architecture is service-oriented mainly service types and their roles inside the scenarios are of main importance in this model. The precise realization of the architecture model leads to implementation specifications for which the implementation model (see 5.2.1.5) is responsible. The implementation model corresponds to the *technology viewpoint* of RM-ODP and basically uses existing specifications for the implementation specifications.

[1] Lead-managed by the Professorship for Geodesy and Geoinformatics at Rostock University.

5.2.1.1 Business model

The business model[2] that corresponds to the *Enterprise Viewpoint* of RM-ODP defines the **objectives**, **aims**, **requirements** and **interests** of the parties concerning the infrastructure. The modeling of business processes is carried out based on selected scenarios, that are reflected as workflows in the process model (see 5.2.1.3). Representative for the multitude of scenarios three are:

- maritime spatial planning
- documentation of expansion efforts of the Elbe river
- MSFD (eutrophication and pollutants)

The *aim* of the MDI-DE is the establishment of a national network for marine data and a geoportal for marine data. The portal should integrate the major data sources ignoring department, government and/or institute boundaries. The infrastructure, which consists of hardware, software, interfaces and organizational *requirements*, is designed to contribute to fulfill current and future reporting and information *obligations* – such as INSPIRE, WFD, MSFD or Natura 2000 – effectively. The MDI-DE establishes a common application platform for distributed systems to *collect and process* marine metadata, data and services. Existing knowledge is made available through its geoportal. Parallel to this development federal and state coastal agencies as well as research institutions have to prepare their own specialized systems so that they can be linked to this infrastructure as infrastructure nodes. The MDI-DE project has an important *coordinating and supporting role* for these preparations. In addition to the technical infrastructure the MDI-DE combines the distributed expertise of the experts on marine and coastal data.

5.2.1.2 Role model

The role model[3] states which actors are inside the system and what roles they have in terms of the diverse tasks that have to be carried out by them. Based on (Dreesmann et al., 2004) and (Rossmanith & Schupp, 2008) the following roles were identified in the MDI-DE:

- *Producer*: Actors generating spatial data by for instance aerial surveys, mapping, sensors and so on. They make these data available to other stakeholders in the form of data collections or use it themselves e.g. to comply with their reporting obligations.
- *Processer*: Actors who use spatial data from other stakeholders and process it (e.g. combine, generalize, extend or refine) in order to generate "new" information. Through this activity this actor automatically belongs to the role *producer*.
- *Broker*: Special role that does not produce or use data but provides services to ensure centralized access to the data of the other actors.
- *Reporting obligated actor*: Actors who are required to follow certain reporting obligations such as the WFD, MSFD and/or INSPIRE and have to provide reports.
- *User*: Generally end users, but also institutions that use spatial data in their own GIS.

Data and services are the main *components* required to establish a SDI. Data is usually categorized in basic geodata, spatial thematic data and metadata. However, since all institutions keep metadata alongside their geodata and only the BSH generates basic geodata this distinction makes little sense in an MDI. Because of that data and services are divided into the following specific characteristics: data (flora [plankton etc.], fauna [fish, sea birds etc.], water quality [chemical measurements etc.], the sea floor) and services (WMS, WFS, CSW, SOS, Gazetteer etc.). An additional component in a marine SDI are the reports at different levels (Europe, Germany, federal states and so on) such as INSPIRE, MSFD, WFD, Habitats Directive and Natura 2000.

By combining the components with the roles different models can be developed for the individual actors. An exemplary model for the Schleswig-Holstein Agency for Coastal Defence, National Park

[2]Comprehensive description can be found in (Korduan & Rüh, 2013, pp. 12).
[3]Comprehensive description can be found in (Korduan & Rüh, 2013, pp. 17).

and Marine Conservation (LKN) is depicted in figure 5.3. The figure shows the roles of the institution and indicates with green colour which data it processes, produces etc. and to what directives it has to report.

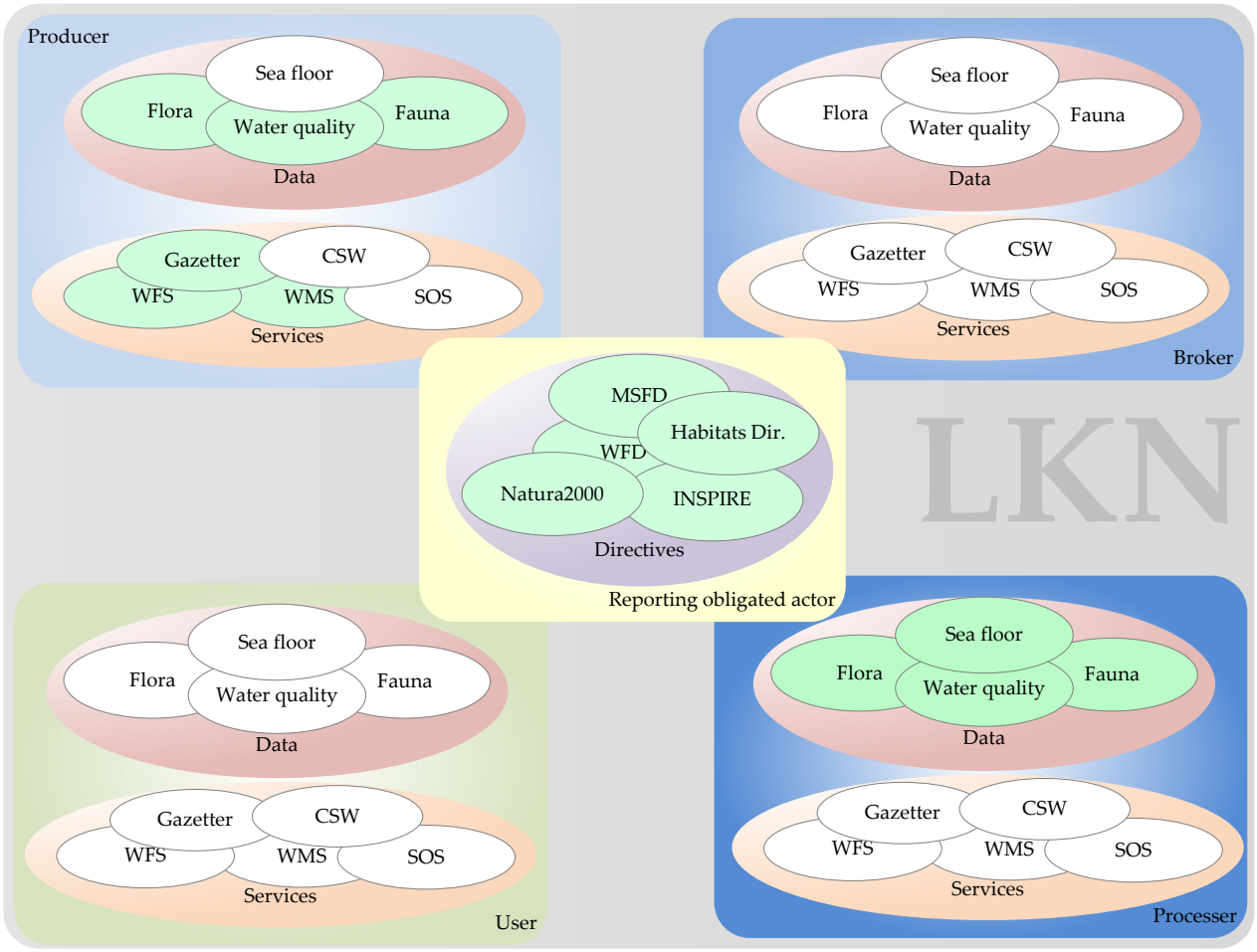


Figure 5.3: *Roles of the actor LKN*

5.2.1.3 Process model

The process model[4] that corresponds to the *engineering viewpoint* describes the activities and data inside the scenarios respectively the processes and models the scenarios with activity- and sequence diagrams. For an exemplary scenario modelled with the help of UML diagrams please refer to subsection 5.2.2 on page 105.

5.2.1.4 Architecture model

The architecture model[5] corresponding to the *information viewpoint* respectively the *computational viewpoint* describes the technical components such as services, interfaces, clients and their responsibilities and inter-relationships with each other. The components of the infrastructure are depicted in figure 5.4. Since the architecture is service-oriented mainly the types of services are stated and what role they play in the interrelations of the scenarios. The specific implementation of the architecture model is carried out in implementation specifications. These are summarized in the implementation

[4]Comprehensive description can be found in (Korduan & Rüh, 2013, pp. 34).
[5]Comprehensive description can be found in (Korduan & Rüh, 2013, pp. 57).

model (see 5.2.1.5). Mainly existing specifications are considered there and will be expanded with profiles to meet the requirements of marine data.

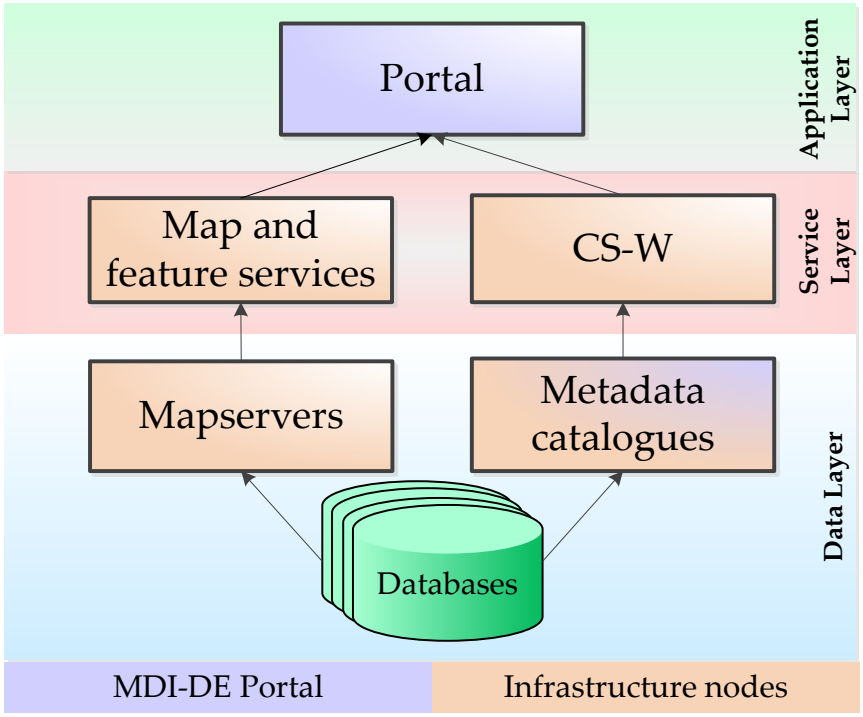


Figure 5.4: *Layered architecture of MDI-DE*

5.2.1.5 Implementation model

The implementation model[6] is based on the architecture model and in that way on the Web Service Architecture of OGC. The *implementation specifications* that are defined here, represent a guide for developers to develop their applications conforming to the OWS architecture. The implementation specifications primarily include rules for XML, HTTP, fixed interface syntax, specific information models for service descriptions and other metadata. Other components defined here are:

- technology decisions,
- development platforms,
- decisions on performance and
- construction and reuse of components.

5.2.2 Exemplary scenario

The *View Model of Software Architecture* (see subsection 2.3.2) underlines the importance of use cases respectively scenarios with its **+1**. This is why the reference model includes a catalogue of scenarios. As an example this section models a scenario with UML diagrams in the MSFD domain, i.e. the workflow enabling the actors to produce the reports the directive presupposes will be modelled. In order to understand this section please refer to subsection 2.5.3 for an overview of the MSFD and the data this directive requires.

How the different submodels of the reference model interact with the connected UML diagrams is outlined by means of an exemplary scenario. In this scenario an actor has to evaluate indicators respectively descriptors to prepare a report to the EU for the MSFD.

The business model defined that the MSFD should have a high priority and the role model

[6]Comprehensive description can be found in (Korduan & Rüh, 2013, pp. 72).

specified that the federal agency for nature conservation (BfN) for instance has to deal with the MSFD in the sense that it records data needed for the reports to the EU. It gathers data about the population of sea birds (descriptor 1 of the MSFD) for example. The different roles of the role model can be recognized in the use-case diagrams. In figure 5.5 the federal agency for nature conservation could be investigator or reviewer as well as both. The process model describes the operations of the scenario, inter alia, with activity diagrams. The activity diagram shown in figure 5.6 takes the use-case "Evaluation of indicators" off the use-case diagram (figure 5.5) and characterizes its course of actions. Firstly an indicator is requested from a service which contains the limits for evaluation. The parameter recorded then gets compared to the limits and if it is under `limit4` but above `limit3` the indicator will be evaluated with *good* and sent back to the service. Figure 5.6 shows as well that the provision of indicators respectively descriptors uses services. This is a link to the architecture model which constituted the orientation towards the usage of services.

5.3 Analysis of existing data sets and services

In the definition by (Groot & McLaughlin, 2000) of the term SDI in section 2.1 on page 16 it was stated that a SDI facilitates *"[. . .] the sharing, access to, and responsible use of geospatial data [. . .] "*. Nowadays there is much geospatial data readily available because agencies have to collect it and have to publish at least some of the data. The reference model (section 5.2) listed all the actors of MDI-DE which have data sets and more importantly services available. This means that when building an SDI these existing data sets will have to be made available through the SDI because SDIs are about the *sharing* and *access to* data as indicated at the beginning. For this reason an overview was needed at the beginning of the project.

To make the existing data sets available the first step is to analyse what already exists in the agencies. To give an overview a web application is created which shows the available data sets and/or services by institution. To fill that application you could ask everyone who might have data sets or services available which would take much time because they potentially would send incomplete information and you would have to ask again. Even if you would send them a form they have to fill out much time would have to be invested to register the information in the web application. Ideally the web application would fetch the currently registered data sets or services dynamically. In order to save time in the registration process and the web application to be dynamic a **database** (schema) is needed (see subsection 5.3.1).

The database enables the registration of data sets and services via web forms (see subsection 5.3.2) which directly store the data in the database which eliminates the time consuming process of filling the database manually. Editing registered data sets or services and storing the modified data in the database also saves time and can be accomplished through the same forms. When some data sets or services are registered it would be helpful to get an overview of what is already registered (see subsection 5.3.3). Using a database this can be done dynamically. This feature is especially helpful for services because these (i.e. its layers) can be displayed directly in the web application.

Summing up the steps to analyse the existing data sets are:

(1) Create a database schema
(2) Allow registration/editing through the use of web forms
(3) Present the already registered services

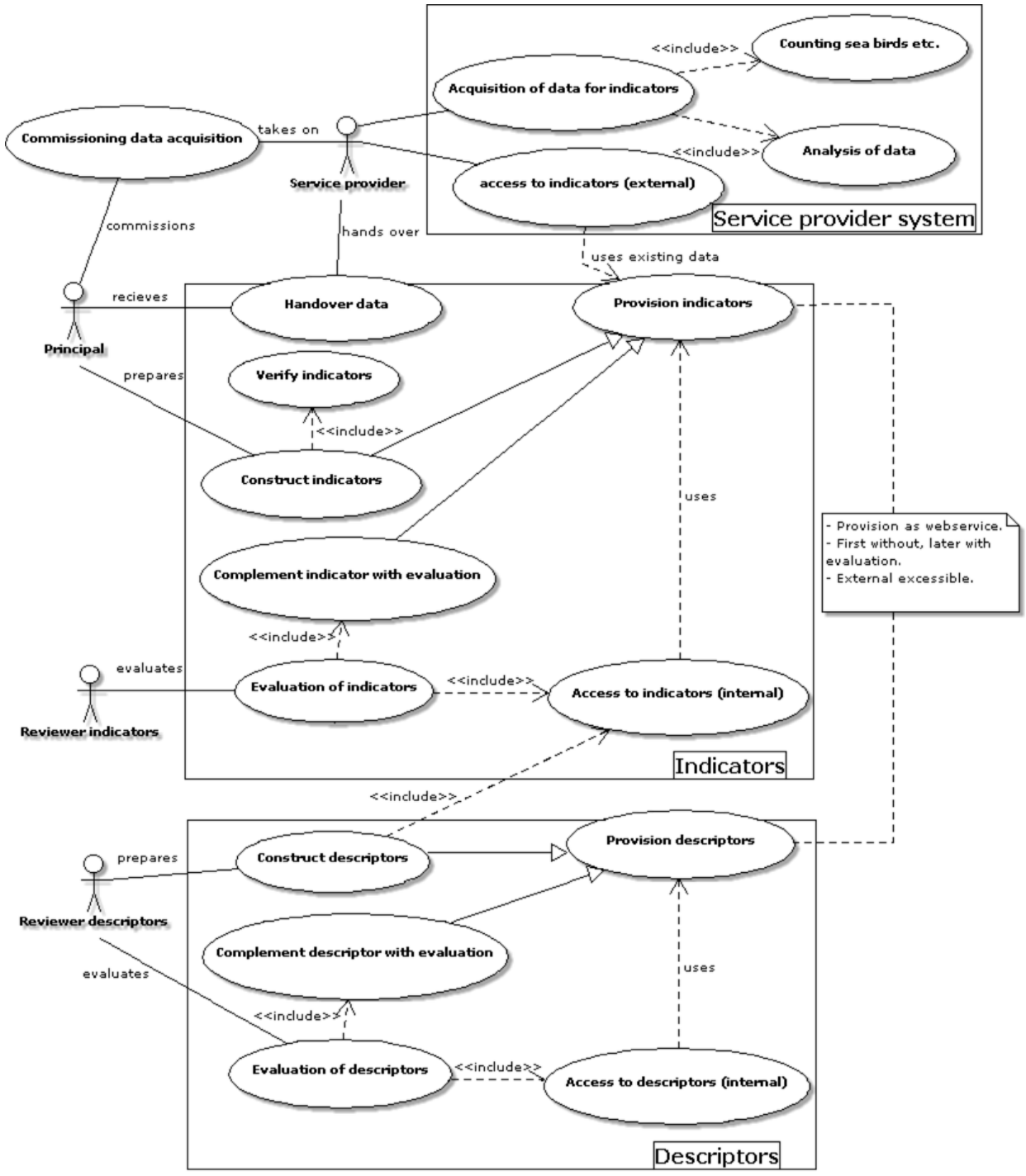


Figure 5.5: *UML use-case diagram to construct respectively evaluate indicators and descriptors*

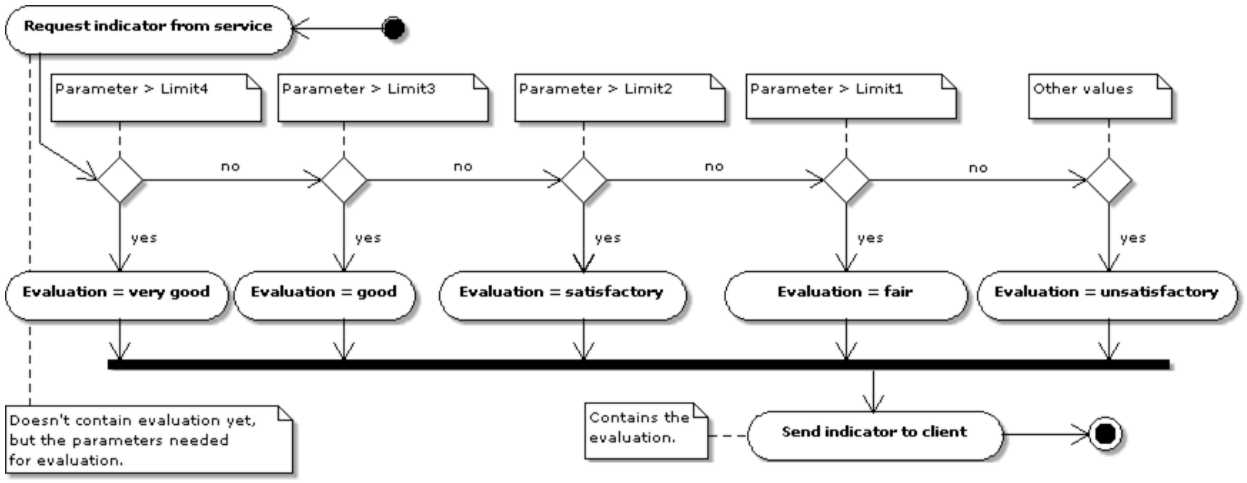


Figure 5.6: *UML-activity diagram to evaluate an indicator*

5.3.1 Creation of a database schema

The database scheme is needed to store all the actors (institutions), their data sets and services as well as the relationships between them and links to EU directives. As depicted in figure 5.7 the three main entities are *Institutions*, *Datasets* and *Services*. The *Institutions* are the central entity because both *Datasets* and *Services* have foreign key relationship to them.

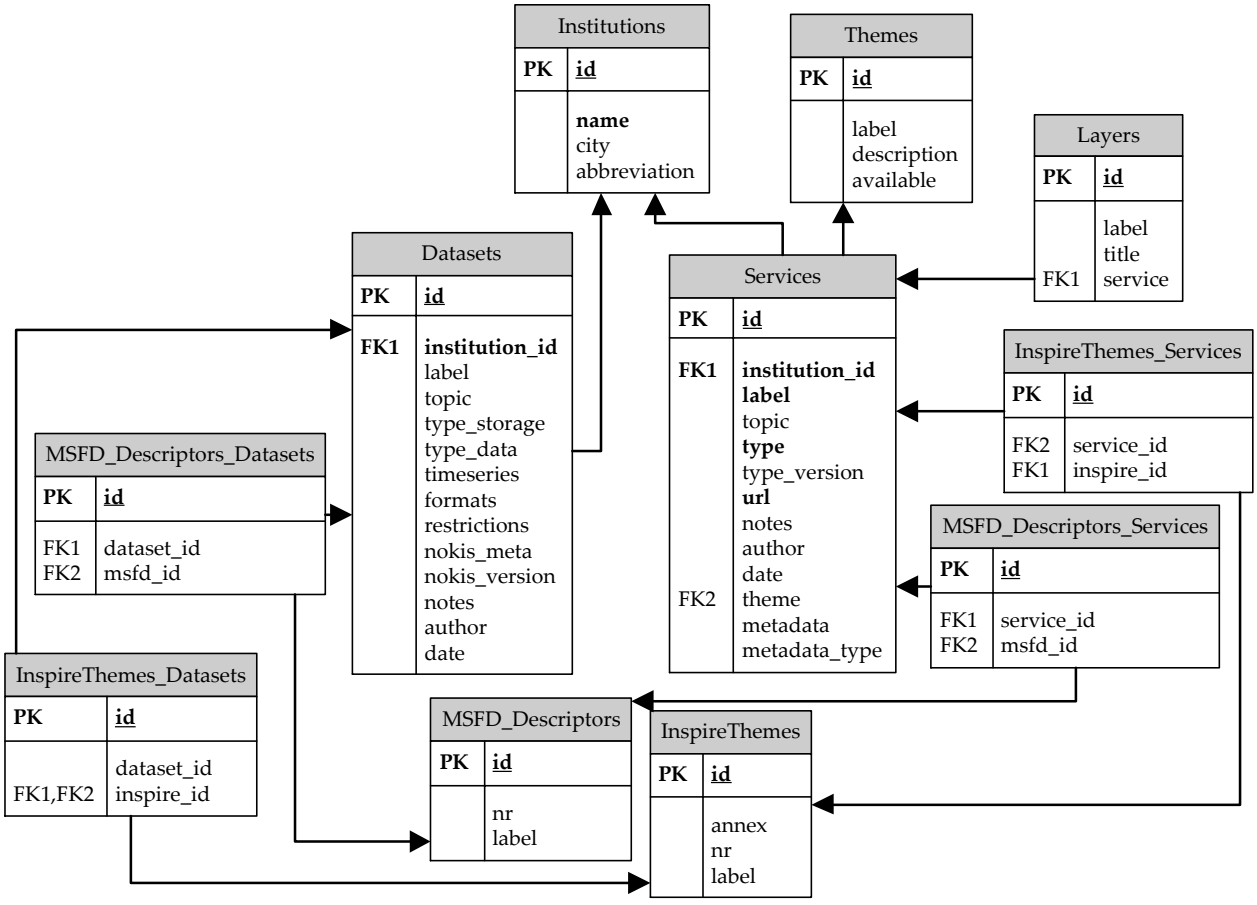


Figure 5.7: *Database scheme to analyse existing data sets or services*

Services have *Layers* which is another entity which stores labels and titles of the layers which are later used for the presentation of the services. Apart from that *Services* have amongst other attributes the attribute `url` which holds the address of the service which is important for presentation purposes. Because a service can deliver data which fulfils INSPIRE or MSFD requirements (see section 2.5 on page 42 for reference) entities to cover this are needed. There is one entity for each directive which stores its descriptors (*MSFD_Descriptors*) respective themes (*InspireThemes*) and one entity for each to store for which theme(s) or descriptor(s) a service provides data (*MSFD_Descriptors_Services* respectively *InspireThemes_Services*).

Datasets are very similar to the *Services* in regard of their relationships to the EU directives INSPIRE (*InspireThemes_Datasets* in place of *InspireThemes_Services*) and MSFD (*MSFD_Descriptors_Datasets* instead of *MSFD_Descriptors_Services*). In addition datasets have attributes which store information on their storage type, format(s), access restriction(s) and more.

To make presentation of the existing data sets and services easier later on, views were defined which for example[7] show (see listing 5.1) what *Services* by certain *Institutions* offer their data for which INSPIRE theme (*InspireThemes* and *InspireThemes_Services*).

```
1  SELECT dienste.id AS dienst_id,
2         inspire_themen.id AS inspire_themen_id,
3         organisationen.name AS instution,
4         dienste.name AS dienst_bez,
5         inspire_themen.annex AS inspire_annex_nr,
6         inspire_themen.nr AS inspire_themen_nr,
7  FROM referencemodel.inspire_themen_dienste,
8         referencemodel.inspire_themen,
9         referencemodel.dienste,
10        referencemodel.organisationen
11 WHERE inspire_themen_dienste.dienst_id = dienste.id
12        AND
13        inspire_themen_dienste.inspire_id = inspire_themen.id;
```

Listing 5.1: *SQL statement to define a view showing services, their institutions and MSFD descriptors they belong to*

5.3.2 Registration of data sets and services

To be able to give an overview over the existing data sets and services they have to be registered and stored in the database first. This was achieved through web forms[8] which used PHP and JavaScript and stored in three files:

(1) PHP form file (see figure 5.8)
(2) CSS file
(3) PHP action file

(1) Apart from giving the service/data set a name, specifying a URL, institution and so on MSFD descriptors and INSPIRE themes can be selected. The styling for labels, legends and form is done with an external CSS file (2). This is also responsible for the delimiter between the institutions (so that they are easier to distinguish/select) and the coloring of the text area currently in focus. Initially, when loading the page three things are fetched from the database: the institutions, MSFD descriptors and INSPIRE themes. If modifying a service the already selected ones are fetched, too. The descriptors/themes fetched from the database are put into a menu of options (HTML `<select>`) displayed on the right side. Figure 5.8 shows an empty menu of options on the left-hand side because nothing has been selected yet. To select or deselect descriptors/themes the buttons << and >> are used. The functionality behind these buttons is delivered through three JavaScript

[7] Another example can be found in listing A.1 on page 145.
[8] Selected excerpts of the accompanying code can be found in listing A.2 on page 145 and listing A.3 on page 147.

functions: `transferOptions()` which uses `addOptions()` and `substractOptions()`. After clicking *send data* the data put in by the user is transferred to (3) via a POST request.

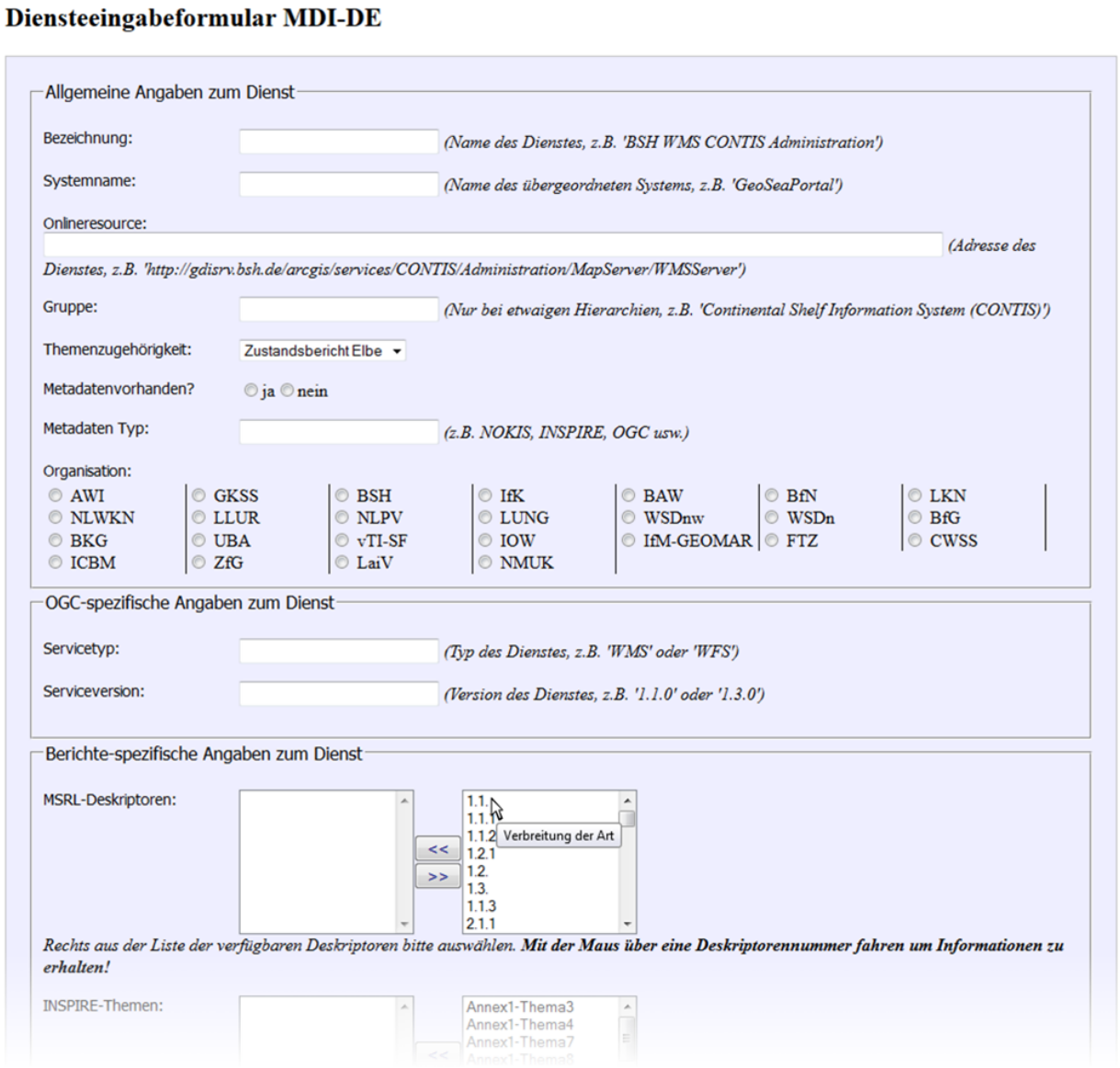


Figure 5.8: *Web form to store an existing service in the database*

If it is a new service or data set the data submitted by the user is just put into the database using a `INSERT INTO` statements. Firstly the "plain" data into the *Services* table and after that the selected descriptors respectively themes into the tables *MSFD_Descriptors_Services* respectively *InspireThemes_-Services*. In case of a new service or data set an `UPDATE` statement is used for the "plain" data and for the selected descriptors respectively themes `DELETE FROM` (firstly all existing relationships are cleared) and `INSERT INTO` themes are used.

5.3.3 Presentation of data sets and services

After creating a database schema and filling the database with existing data sets respectively services an overview of what has already been submitted can be given now. Apart from a simple page with a table for data sets respectively services an advanced page was developed for services offering the functions shown in figure 5.9[9].

[9] see code excerpt A.4 on page 149

The advanced page itself shows all registered services (their titles) assigned to its providing institution. Each service can be expanded to see all its layers. The layers have an option box in front of them which is used for the *WMSBrowser Preview combined* functionality explained later. The page also contains the links to the functionalities shown in figure 5.9. ***OpenLayers Preview*** When clicking on the title of a service the user is led to a page which offers a preview of all the layers of the service. Which service was selected is determined through a request parameter in the URL (`id`) which is read by a HTTP GET request. For the visualization an OpenLayers interface and OpenStreetMap (OSM) or Demis as base maps are used. OpenLayers is a free, Open Source JavaScript library that – according to its homepage (openlayers.org) – *"[. . .] makes it easy to put a dynamic map in any web page."* The layer tree is a widget provided by MapFish (inter alia an OpenLayers JavaScript toolkit, mapfish.org) to control OpenLayers layers because such functionality is not implemented very well in OpenLayers. ***WMSBrowser Preview*** Clicking on *View in WMSBrowser* offers another user interface showing the layers of a service using a different technology but the same passing of the parameter `id` of the service[10]. WMSBrowser is an "add-on" of GeoExt (in GeoExt terms a *GeoExt ux*[11]). GeoExt – according to its website (www.geoext.org) – is a *"JavaScript Toolkit for Rich Web Mapping Applications"* which *" [. . .] brings together the geospatial know how of OpenLayers with the user interface savvy of Ext JS [. . .] "*. This points out that GeoExt uses components of Ext JS and OpenLayers. In the developed application inter alia a simple data store and tree panel is used by Ext JS and a map and several layer objects are provided by OpenLayers. ***WMSBrowser Preview combined*** As stated in the introduction of the advanced page itself the layers of the services can be expanded and with the option box in front of them can be selected. This enables a combined view with several layers from different services in one map using the WMSBrowser again. However, now a HTTP POST request is used to pass the selected layers because it is not just one service anymore. ***Capabilities*** When clicking on *Capabilities* the user is directed to the server and its capabilities in pure XML format. Because this is hard to read for users and can be quite lengthy a user can also click on *Capabilities in human-readable format*. This uses the `CapabilitiesParser` by Tobin Bradley[12] and shows a summary of the service (name, version, supported CRS, list of layers).

Because it would not make sense to let the users type in the titles and names of the layers of the services which are needed for presentation there is a functionality at the bottom of the advanced page. This PHP script empties the *layers* table in the database and iterates through the list of services sending a `GetCapabilities` to each. The resulting XML is held in a PHP `DOM Document` object from which information about the layers is extracted and then stored in the database.

5.4 Evaluation of MDI-DE services

After the existing services are known (see section 5.3) and after additional ones were set up based on existing data sets an indispensable prerequisite for an interoperable architecture are useful and functional services. This is why this section evaluates the services currently available on the MDI-DE geoportal[13]. Section 5.3 outlined how existing services can be made to be known by people. Furthermore services are useful only if they conform to certain standards, are available most of the time and deliver a good performance. Conforming to standards means that it is always clear what kind of requests the service understands and what its response will look like so that systems are able to work together. This will be examined in subsection 5.4.1. Availability means that a service is up and running and responding to requests most of the time (for which varying definitions exists). Good performance means that users will be satisfied by the turnaround time of a response. Both quality of service issues will be addressed in subsection 5.4.2.

[10] see listing A.5 on page 151
[11] http://trac.geoext.org/wiki/ux
[12] http://code.google.com/p/metadata-navigator/source/browse/services/wms-parser.php
[13] As of mid-2013.

Figure 5.9: *Web page depicting services and their layers and advanced functions*

Section 3.6 introduced several tools to test performance and conformity. This section will mainly use free and open-source software (FOSS) because there simply were no funds for commercial tools. The exception is the *sdi.suite serviceMonitor* which is already included in the sdi.suite used in the MDI-DE project.

For both performance and conformity all the tools were used that were introduced in section 3.6 with the exception of *geOps MapMatters* that was not usable due to lacking service statistics.

Load tools introduced in 3.6.3.4 where not used because firstly it is not known how many users are using the service(s) in the first place. This means that generating simultaneous requests does not make sense because it cannot be said how many requests there really were in total. Secondly producing load means increased bandwidth etc. which result in higher costs which can be a problem for the institutions involved.

5.4.1 Conformity with INSPIRE and OGC

This section evaluates if or to what extent the MDI-DE services conform to OGC and INSPIRE requirements. While the OGC TEAM engine (5.4.1.1) checks for OGC conformity the Neogeo WMS INSPIRE tester (5.4.1.2) and INSPIRE Geoportal Metadata Validator (5.4.1.3) check for INSPIRE conformity. The GDI-DE Testsuite is capable of evaluating both which is done in 5.4.1.4. At the end of this section the results will be discussed and compared (5.4.1.5).

5.4.1.1 OGC TEAM engine

OGC's TEAM engine tests services against the OGC specifications as 3.6.2.1 on page 69 pointed out. To test web map services (WMS) there are test suites available for version 1.1.1 and 1.3.0 which cover all MDI-DE services. However, for web feature services there are only test suites available for version 1.0.0 and 1.1.0 but all WFS's inside MDI-DE are version 2.0.0 WFS. This is why there are no results for WFS.

The results for WMS had to be put into perspective because the number of tests which a service has to pass varies largely. One service had to pass 625 tests while the one with the fewest tests

only had to pass 55. The number of failed tests is the most important thing to know about a service but does not help to compare services in this case. For example the services "Kartenportal Umwelt MV" and "Schutzgebiete" both by the Agency for the Environment, Nature Conservation and Geology of Mecklenburg-Western Pomerania (LUNG) have failed ten respectively nine tests. However, "Kartenportal Umwelt MV" passed 465 tests which means the ratio between passed and failed tests is 2,11 % while this ratio is 13,85 % for "Schutzgebiete" because it passed just 56. This is why the pure numbers do not really help to get a quick glance at the results. Figure 5.10 depicts a diagram with the percentage of passed compared to failed tests making comparing the services with each other easy. The reason for the varying number of tests lies in the version, number of layers, CRS, accepted formats and so on.

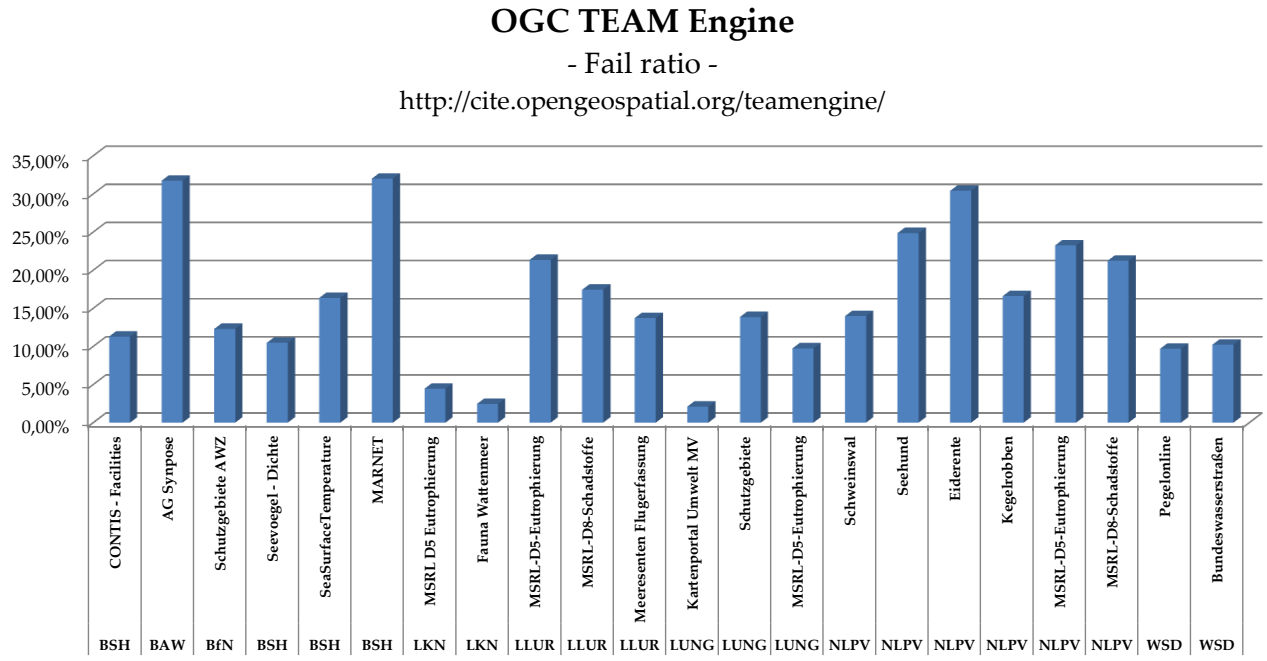


Figure 5.10: *OGC conformity of MDI-DE services (OGC TEAM engine)*

WMS 1.1.1 often fail tests because of the various formats that should be supported when an exception is raised. For instance that XML should be the default format (1) or that only red pixels should be returned (2). However, in at least two cases but probably more these result were false negative errors because the service returned XML as default (1) and another service returned red pixels (2). Unfortunately the same thing happens now and then with the format of a GetCapabilities request. Another common reason which results in a few failed tests is the GetFeatureInfo request or more precisely the formats of the response. Because if the layer requested is not queryable it fails the tests to respond to a GetFeatureInfo request with plain text, HTML and GML.

A common reason for a **WMS 1.3.0** to fail tests is that it does not return transparent pixels only when a GetMap request is made with `TRANSPARENT=TRUE` and an "empty" bounding box. Another common problem is the usage of exponential notation values for the bounding box parameter in a GetMap request where a service fails the test when it returns an invalid response. Furthermore when the size of the LegendGraphic is not exactly 20x20 pixels a test is failed. Unfortunately, when a service offers many layers in many different styles this behavior results in a vast amount of failed tests.

5.4.1.2 Neogeo WMS INSPIRE tester

Section 3.6.2.2 on page 69 stated that *Neogeo WMS INSPIRE tester* is used to test services regarding INSPIRE conformity. The application makes use of the numbering of the requirements and recom-

mendations of INSPIRE using `TG_Req#XX` for requirements and `TG_Rec#XX` for recommendations. As the name indicates it can only be used for WMS which is why the few WFS services had to be left out in the evaluation.

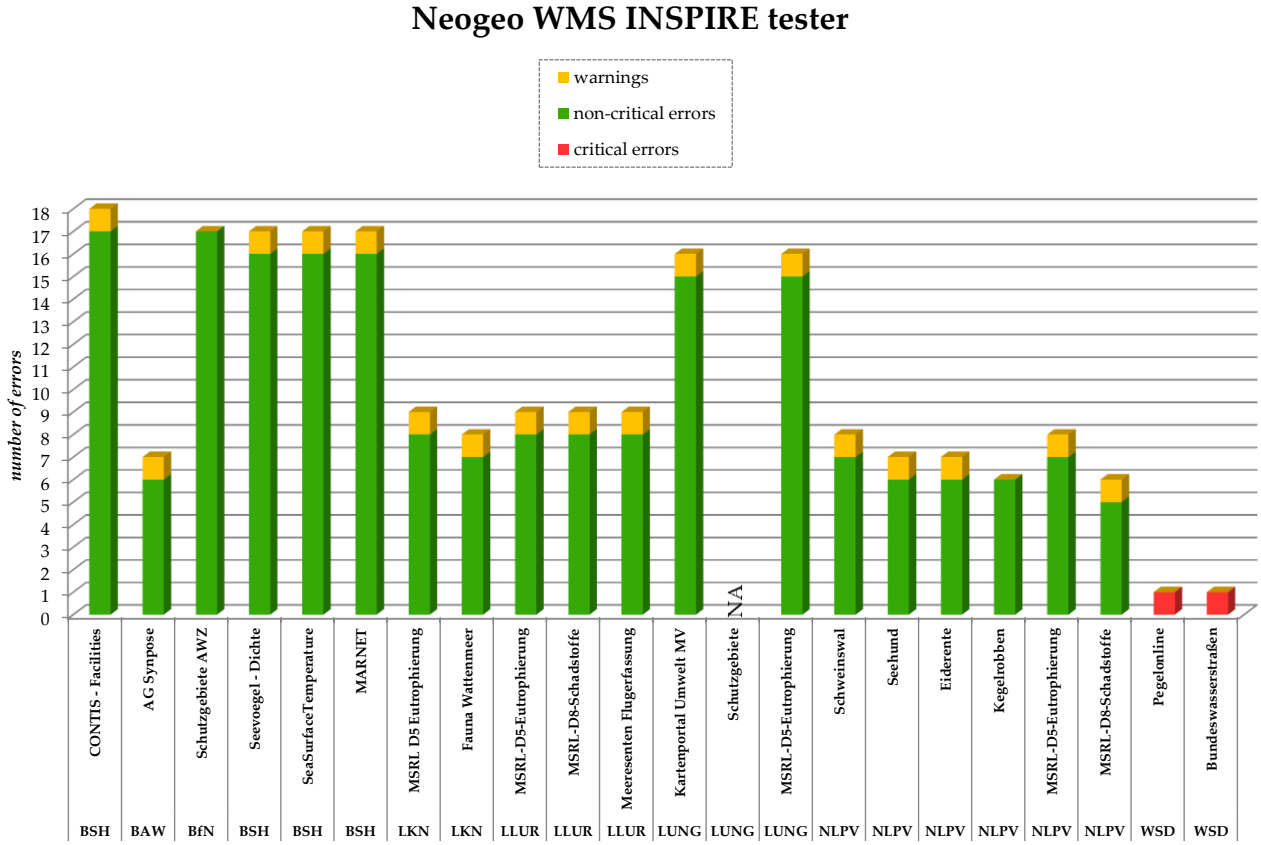


Figure 5.11: *INSPIRE conformity of MDI-DE services (Neogeo WMS INSPIRE tester)*

Figure 5.11 depicts how the MDI-DE services perform. However, two services seem to have a critical error. But the reason for that is that these two services are WMS 1.1.1 and not 1.3.0 as INSPIRE requires. When these service respectively their servers will be updated to 1.3.0 fortunately there will be no services failing INSPIRE conformity because of critical errors. Common non-critical errors include inter alia

- *"BoundingBox missing for some CRS - TG_Req#36"*
- *"The element <inspire_common:MetadataUrl> is not present in the element <inspire_vs:ExtendedCapabilities> - TG_Req#06"*
- *"The <wms:Fees> element is mandatory. This element was not found. - TG_Req#10 and TG_Req#124"*
- *"The default language shall be declared in the <inspire_common:DefaultLanguage> element. - TG_Req#71"*
- and others concerning language (*ResponseLanguage*, *SupportedLanguages* etc.)

This means although some services have up to 17 non-critical errors this is not a huge problem because the services can be changed rather easily to get fully INSPIRE complaint. The only warning found was *"Limitations on public access - <wms:AccessConstraints> - unexpected value - TG_Rec#05"*. This is because the services used `keine` although INSPIRE recommends in its technical guideline (Craglia, 2010a) to use the codelist B.5.24 of ISO 19115 which means that only `copyright`, `patent`, `patentPending`, `trademark`, `license`, `intellectualPropertyRights`, `restricted` and `otherRestrictions` would be allowed. Figure 5.11 also shows that there is no data available about the service "Schutzgebiete" by the Agency for the Environment, Nature Conservation and Geology of Mecklenburg-Western Pomerania (LUNG). It remains unknown why the tool does not

accept this service (*" The provided URL doesn't seem to match a WMS service. We can't proceed to any further test on this URL"*).

5.4.1.3 INSPIRE Geoportal Metadata Validator

Section 3.6.2.3 on page 69 introduced the INSPIRE Geoportal Metadata Validator with which WFS are also testable (since this is all about metadata). This tool yields *Schema Validation Issues* and *Inspire Validation Issues* but because Schema Validation Issues are most *" [. . .] likely [. . .] at the origin of the issues reported in the 'Inspire Validation Issues [. . .] "'* (from the report) only these issues will be analysed. As figure 5.12 depicts the issues reach from just one to sixteen.

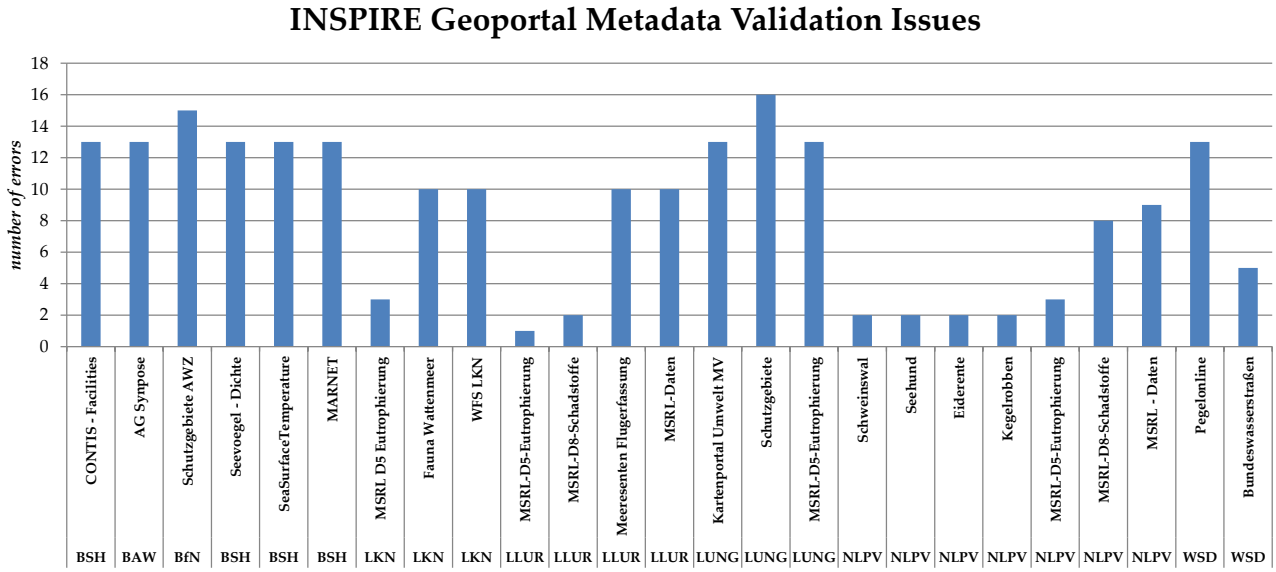


Figure 5.12: *INSPIRE conformity of MDI-DE services (INSPIRE Geoportal Metadata Validator)*

The most common validation issues almost every service had were

- *the INSPIRE Extended Capabilities element could not be read*,
- *one or more layers failed the INSPIRE validation* and
- *the metadata element "Layers" is missing, empty or incomplete but it is required*.

Other very common validation issues of services with X or more issues include inter alia

- *the metadata element "Responsible Organisation" is missing, empty or incomplete but it is required*,
- *the metadata element "Temporal Reference (at least one among Date of Creation, Date Of Publication, Date Of Last Revision, Temporal Extent)" is missing, empty or incomplete but it is required*,
- *the metadata element "Resource Type" is missing, empty or incomplete but it is required*,
- *the metadata element "Metadata Point Of Contact" is missing, empty or incomplete but it is required*,
- *the metadata element "Metadata Date" is missing, empty or incomplete but it is required*,
- *the metadata element "Metadata Language" is missing, empty or incomplete but it is required*,
- *the metadata element "Spatial Data Service Type" is missing, empty or incomplete but it is required*,
- *the metadata element "Response Language" is missing, empty or incomplete but it is required*,
- *the metadata element "Supported Languages" is missing, empty or incomplete but it is required*,
- *the metadata element "Limitations On Public Access" is missing, empty or incomplete but it is required*,
- *the metadata element "Conditions For Access And Use" is missing, empty or incomplete but it is required*,
- *the metadata element "Mandatory Keyword" is missing, empty or incomplete but it is required* and
- *the metadata element "Resource Abstract" is missing, empty or incomplete but it is required*.

Rare validation issues were

- *the Service Metadata URL "http://<URL>" is invalid because: "No resource could be found."* and
- *the network service contacted with " http://<URL>" declares the default language ("eng") which is not the one used for the response ("ger").*

5.4.1.4 GDI-DE Testsuite

Subsection 3.6.4 on page 75 stated that the GDI-DE Testsuite can not only be used to check conformity with INSPIRE and OGC requirements but also to check quality of service. While this will be discussed in 5.4.2.3 on page 124 the conformity checks will be discussed in this section.

Not all test classes were used inter alia because there were no catalogue services available and the INSPIRE metadata tests are performed within the *INSPIRE View Services* test class anyway. The test classes used are:

(1) INSPIRE View Services based on WMS 1.1.1/1.3.0 (conformity to INSPIRE View Service requirements)
   - Conformity classes
     (a) Mandatory: INSPIRE Requirements Network Services
     (b) Optional: INSPIRE Requirements Interoperability

(2) GDI-DE WMS 1.1.1/1.3.0 (conformity to OGC Web Map Server Implementation Specification 1.1.1/1.3.0)
   - Conformity classes
     (a) Mandatory: ISO Requirements
     (b) Optional: ISO Optional

OGC conformity

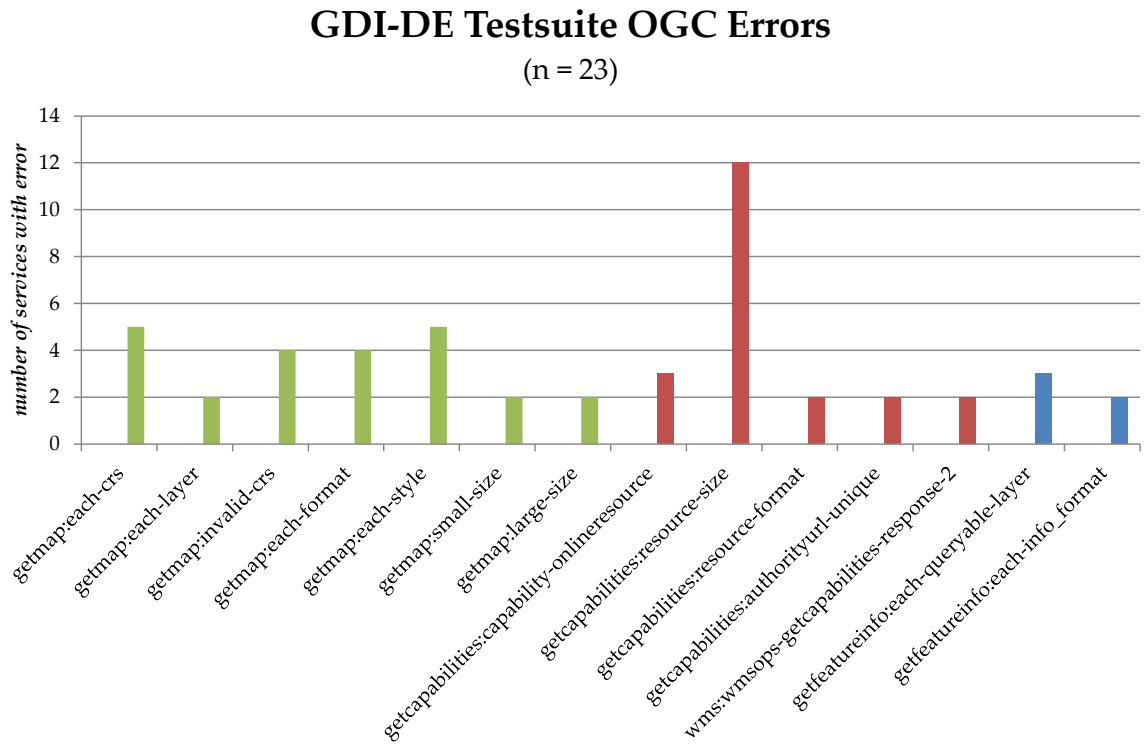


Figure 5.13: *OGC conformity of MDI-DE services (GDI-DE Testsuite)*

Although only one service passed OGC conformity and another one passed with warnings overall the OGC conformity is good. 19 of the 23 tested services showed four or less errors and only two had more than four errors. Figure 5.13 depicts the common errors the services showed. Common means that at least two services were found with a specific error. The errors were:

- GetMap
  - Invalid response when requesting a layer with a certain CRS (which the service states as supported, `getmap:each-crs`)
  - Invalid response when requesting a layer with the LAYERS parameter set to a specific layer (`getmap:each-layer`)
  - No service exception (code=InvalidCRS) when requesting a layer with the CRS parameter set to an invalid CRS (`getmap:invalid-crs`)
  - MIME type of the response matches the format image/png when the FORMAT parameter is set to image/png (`getmap:each-format`)
  - Invalid MIME type of response when a specific style is requested with the STYLES parameter (`getmap:each-style`)
  - Returned image with WIDTH=8 and HEIGHT=5 is not exactly that size (`getmap:small-size`)
  - Returned image for a large map (1024x768 or largest map supported) is not exactly that size (`getmap:large-size`)
- GetCapabilities
  - Not every OnlineResource URL intended for HTTP Get requests in the capabilities document is a URL prefix (`getcapabilities:capabilityonlineresource`)
  - The size of the LegendURL for Style default in a layer is not 100x512 (`getcapabilities:-resource-size`)
  - MIME-type returned for the MetadataURL for a specific layer is text/plain (`getcapabilities:resource-format`)
  - There are no child Layer elements in the capabilities document that define an AuthorityURL with the same name attribute as one inherited from a parent layer (`getcapabilities:-authorityurl-unique`)
  - The response references a valid and accessible copy of the DTD in Annex A.1 and validates with it (`wms:wmsops-getcapabilities-response2`)
- GetFeatureInfo
  - Invalid MIME type of response when requesting a layer with QUERY_LAYERS parameter set to a specific layer (`getfeatureinfo:each-queryable-layer`)
  - MIME type of the response does not match the format specified by the INFO_FORMAT parameter (`getfeatureinfo:each-info_format`)

The service *Kartenportal Umwelt MV* by the LUNG had many additional unique errors while the services *Schutzgebiete* (also by the LUNG) and *MSRL-D5-Eutrophierung* (by the NLPV) had two additional unique errors. The only warning found in four services was that a response should still be valid when a GetFeatureInfo request contains an undefined parameter (`basic_elements:extra-GetFeatureInfo-param`).

INSPIRE conformity

To test for INSPIRE conformity both conformity classes (mandatory: INSPIRE Requirements Network Services and optional: INSPIRE Requirements Interoperability) were combined. Apart from one service which passed with warnings the results are split into two groups. There is one group with ten services which had up to five errors and another group with nine services which showed eight or more errors. How common (or rare) the different errors were is depicted in figure 5.14.

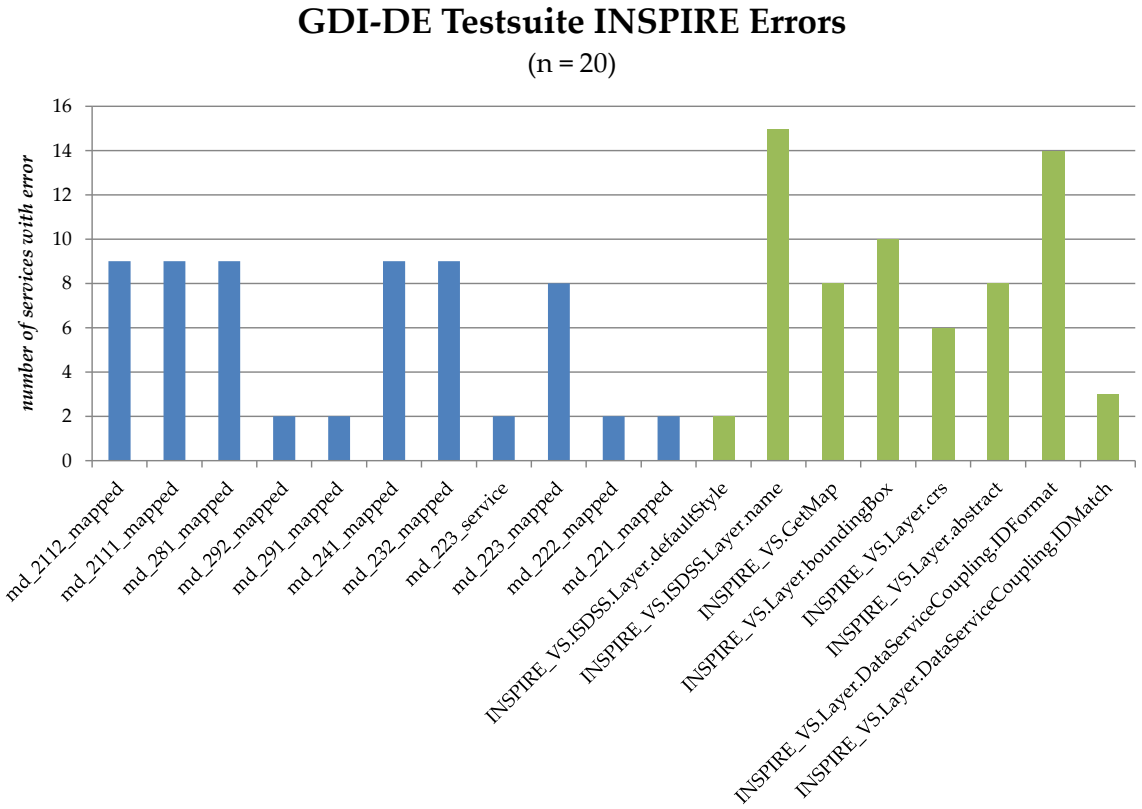


Figure 5.14: *OGC conformity of MDI-DE services (GDI-DE Testsuite)*

The errors (see figure 5.14) are in detail:

- Metadata
  - A metadata date is not given (`md_2112_mapped`)
  - No point of contact specified (`md_2111_mapped`)
  - A conformity statement with a result of conformance evaluation ist not given (`md_281_mapped`)
  - No conditions applying to access and use are described for the resource at least once (`md_292_mapped`)
  - Limitations on public access must be described at least once for the resource (`md_291_mapped`)
  - No mandatory keyword describing the service specified (which must be one of the keywords described in the INSPIRE directive, `md_241_mapped`)
  - Spatial data service type (set to "view") not specified (`md_232_mapped`)
  - Resource type ("service") not specified (`md_223_mapped`)
  - No keyword from the INSPIRE directive specified (`md_223_service`)
  - Resource abstract which describes the resource not specified (`md_222_mapped`)
  - Service title for identification not specified (`md_221_mapped`)
- Network Service
  - Default language missing (`INSPIRE_VS.GetMap`)
  - Not every layer has a default style (`INSPIRE_VS.ISDSS.Layer.defaultStyle`)
  - Not every layer has a name (`INSPIRE_VS.ISDSS.Layer.name`)
  - Not every layer has defined a Bounding Box for every CRS (`INSPIRE_VS.Layer.boundingBox`)
  - Not every layer supports EPSG:4326 and EPSG::4258 (`INSPIRE_VS.Layer.crs`)
  - Not every layer has an abstract (`INSPIRE_VS.Layer.abstract`)
  - Identifiers coupling a layer with its metadata are absent or not identical (`INSPIRE_VS.Layer.DataServiceCoupling.IDMatch`)

- Not every layer has an AuthorityURL and a correctly formatted identifier (`INSPIRE_VS.Layer.DataServiceCoupling.IDFormat`)

There were three warnings regarding metadata and three regarding the service itself. However the three concerning the service itself are confusing and were already included in the errors (`INSPIRE_VS.Layer.crs` and existance of an abstract for the service and a style for every layer) and it is unkown how it is possible for them to be just a warning because e.g. either there is an abstract or there is none. The three "real" warnings were:

- No citation of the product specification or user requirement against which data is being evaluated given (`md_282_mapped`)
- The GML version could not be determined (`md_start`)
- Resource locator not provided for a linkage (`md_224`)

Discussion

Although the GDI-DE Testsuite is quite advanced and offers many functions there were two main issues. **Firstly** it offers an API to access the results (conformity and QoS). However, the API lacks a documentation completely but it is described by a Web Services Description Language (WSDL) document. This makes the API useable though as listing 5.2 shows, but unfortunately the results are useless for both conformity and QoS. The support of the GDI-DE Testsuite was not able to help with the issues either.

Secondly it was not possible to evaluate some services because there was an error message suggesting that the URL was invalid. However, these services worked well accessing them with other tools.

```php
<?php
  //$testConfID = (int) 6193; //[Name] => BfN INSPIRE WMS
  $testConfID = (int) 14048; //LKN Fauna Wattenmeer WMS
  //period = enum -> single, hour, day, week, month, year
  $period = "week";
  $userName = "User";
  $pass = "Password";
  $client = new SoapClient("http://testsuite.gdi-de.org/gdi/download?id=wsdl");

  //<<getTestReport>>
  //getTestConfigurations() to get the testConfID
  $testConfigurations = $client->getTestConfigurations($userName,$pass);

  //startTestConfiguration(userName pass testConfID) to get reportID
  $reportID = $client->startTestConfiguration($userName,$pass,$testConfID);

  $result=$client->getTestReport($userName,$pass,$reportID);
  //getTestReportXML or getTestReportPDF also possible

  //<<QoS>>
  $getQOSConfigurations = $client->getQOSConfigurations($userName,$pass);

  //getQoSTestResult(userName pass TestConfID Period)
  $QoSTestResult = $client->getQoSTestResult($userName,$pass,$testConfID,$period);
?>
```

Listing 5.2: *Basic code to access the GDI-DE Testsuite API*

5.4.1.5 Discussion and comparison of the results

OGC

As 5.4.1.1 pointed out OGC's TEAM Engine produced several false negatives which influences the number of "real" errors. Another factor influencing the number of errors is the number of layers a service offers. In conjunction with minor errors like *LegendGraphic is not exactly 20x20 pixels* (bigger LegendGraphics were a requirement in the MDI-DE project) this results in a large number of errors. Thus the expressiveness of figure 5.10 on page 113 is rather small and a case-by-case review is needed for each service.

Interestingly as table 5.1 shows the results of OGC's TEAM Engine and the GDI-DE Testsuite only share one error. This means that both tools are needed to detect all errors.

Table 5.1: *OGC error comparison*

No	Error	TEAM Engine	Testsuite
01	Wrong format when exception is raised	+	-
02	GetFeatureInfo request to non-queryable services	+	-
03	GetMap request with `TRANSPARENT=TRUE` and "empty" bounding box	+	-
04	Use of exponential notation values for the bounding box	+	-
05	LegendGraphic is not exactly _ pixels	20x20	100x512
06	Invalid response when requesting a layer with a certain CRS	-	+
07	Invalid response when requesting a layer with the LAYERS parameter set to a specific layer	-	+
08	No service exception when requesting a layer with the CRS parameter set to an invalid CRS	-	+
09	MIME type of the response matches the format image/png when the FORMAT parameter is set to image/png	-	+
10	...		

INSPIRE

Table 5.2 shows that the *Neogeo WMS INSPIRE tester* detects the fewest errors of the three INSPIRE conformance test tools. The *INSPIRE Geoportal Metadata Validator* finds almost as many errors as the *GDI-DE Testsuite* does. However, the GDI-DE Testsuite not only detects more errors furthermore it is more precise on the errors (e.g. the error *metadata element "Layers" is missing, empty or incomplete but it is required* from INSPIRE Geoportal is clearer in the GDI-DE Testsuite because it details what is missing respectively why it is incomplete). But since not all services could be tested with the GDI-DE Testsuite it cannot be favoured over the INSPIRE Geoportal Metadata Validator.

5.4.2 Performance and availability

Quality of service is an important measure especially for user satisfaction. The tools FGDC Service Status Checker (SSC) (5.4.2.1), sdi.suite serviceMonitor (5.4.2.2) and GDI-DE Testsuite (5.4.2.3) were used to monitor the services for a specific time period. The results of the measurements are discussed and compared in 5.4.2.4. 5.4.2.4 highlights issues with measurements respectively monitoring in general, too.

Table 5.2: *INSPIRE error comparison*[a]

No	Error	N	I	G
01	Unexpected value for AccessConstraints	+	+	+
02	BoundingBox missing for some CRS	+	-	+
03	MetadataUrl not found in ExtendedCapabilities	+	+	+
04	Fees element mandatory but not found	+	+	-
05	DefaultLanguage, ResponseLanguage, . . .	+	+	+
06	One or more layers failed the INSPIRE validation	-	+	+
07	Metadata element ”Layers” is missing, empty or incomplete but it is required	-	+	+
08	Metadata element ”Responsible Organisation” is missing, . . .	-	+	-
09	Metadata element ”Temporal Reference” is missing, . . .	-	+	+
10	Metadata element ”Resource Type” is missing, . . .	-	+	+
11	Metadata element ”Metadata Point Of Contact” missing, . . .	-	+	+
12	Metadata element ”Metadata Date” is missing, . . .	-	+	+
13	Metadata element ”Metadata Language” is missing, . . .	-	+	+
14	Metadata element ”Spatial Data Service Type” is missing, . . .	-	+	+
15	Metadata element ”Mandatory Keyword” is missing, . . .	-	+	+
16	Metadata element ”Resource Abstract” is missing, . . .	-	+	+
17	Not every layer supports EPSG:4326 and EPSG::4258	-	-	+
18	Not every layer has an AuthorityURL and a correctly formatted identifier	-	-	+
19	Service title for identification not specified	-	-	+
20	A conformity statement with a result of conformance evaluation ist not given	-	-	+

[a]N = Neogeo WMS INSPIRE tester, I = INSPIRE Geoportal Metadata Validator, G = GDI-DE Testsuite

5.4.2.1 FGDC Service Status Checker (SSC)

Because the API of the FGDC Service Status Checker returns 1 000 records at most and there were over 3 000 records since registration the time range had to be reduced from mid-August to beginning of September 2013 (which resulted in 984 records). As depicted in figure 5.15 the results for the scores is split in two groups. The one group (eleven services) achieved scores above 95 % while the other group (nine services, excluding the “NA” [not available] ones) got scores under 50 %. Since the documentation[14] states that the score*“[. . .] is calculated as a factor of the Speed performance and the correctness of the response”* a look at the speed performance (response time) helps to understand the score results. Unfortunately – as shown in figure 5.16 – the speed just adds to the confusion as services with high (=bad) response times have a very high score while other services with very low (=good) response times have very low scores. Thus their responses must be very incorrect. However, this cannot be told using SSC. Looking at the considerations regarding OGC conformity from the previous subsection the low scores still do not add up (especially for the services by the LKN). Furthermore it is unknown why some tests could not be performed on specific services (the ones having “NA” [not available] instead of a bar in the figures).

[14]http://registry.fgdc.gov/statuschecker/documentation.php#service-scoring

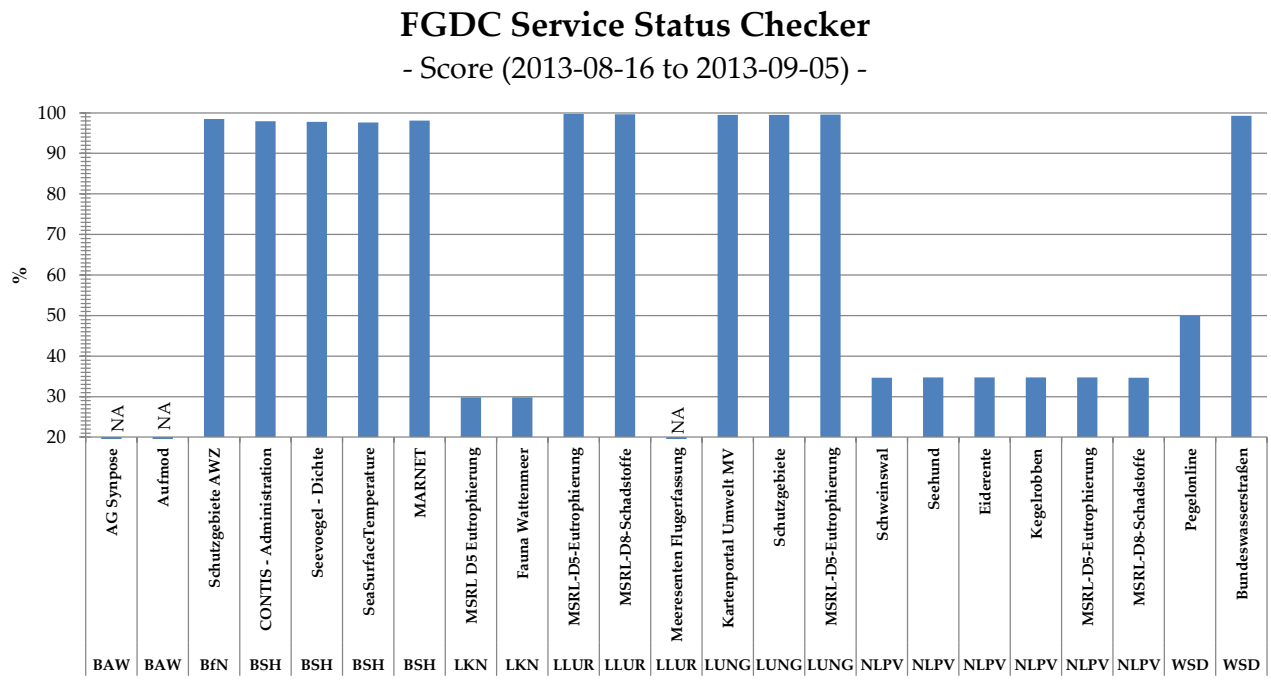


Figure 5.15: *Results for score measured by FGDC Service Status Checker*

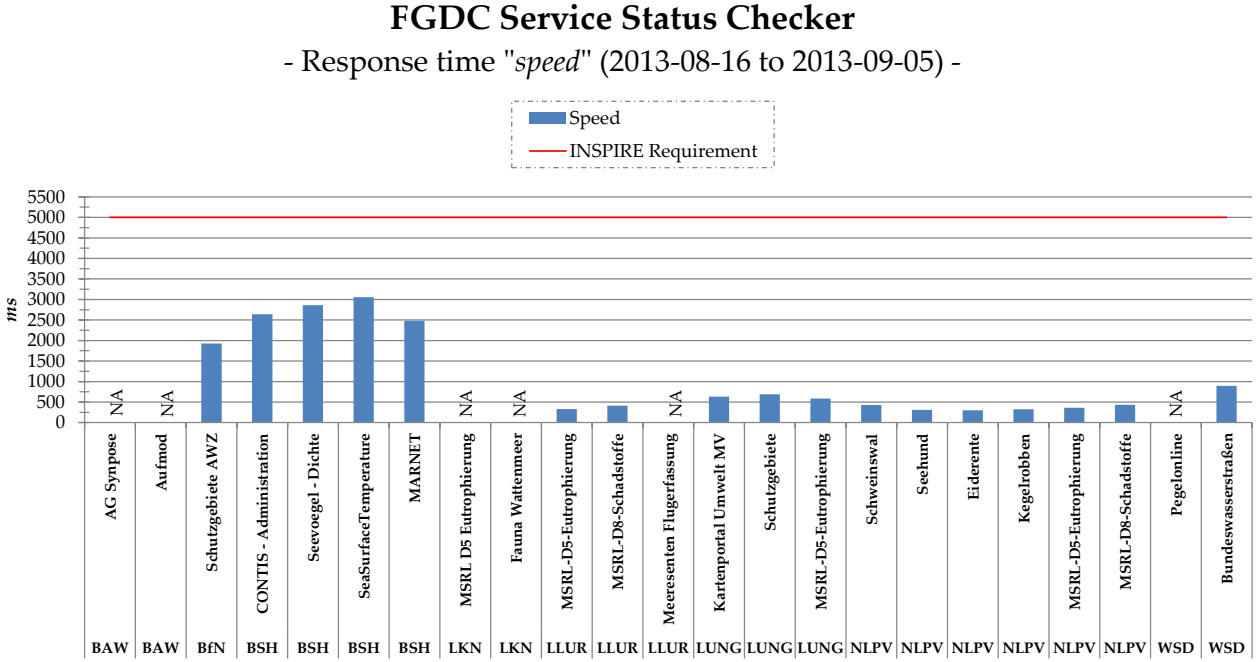


Figure 5.16: *Results for speed measured by FGDC Service Status Checker*

5.4.2.2 sdi.suite serviceMonitor

The sdi.suite serviceMonitor offers four parameters regarding quality of service and measurements for different time spans (last 30 days, since registration and so on). The four parameters are:

(1) Availability
(2) Response time
(3) Service Exceptions
(4) HTTP Errors

In this study only (1) and (2) will be used because these are the most important parameters comparable with the other tools. The measurements with the sdi.suite serviceMonitor were carried out over a period of roughly four months (April to August 2013). As figure 5.17 shows the response times differ for GetCapabilities and GetMap/GetFeature request but there is no indication that either one is slower or faster generally. The figure also shows that the response time fulfils the INSPIRE requirement of 5 000 ms. Figure 5.18 shows the availability of all services is below the INSPIRE requirement of 99 %. Interestingly the availability for the different request types is not the same as the diagram shows. Both the response times as well as the availability results are unevenly distributed for the different requests. This means that sometimes the response time for a GetCapabilities request is shorter than for a GetMap request and sometimes the other way round.

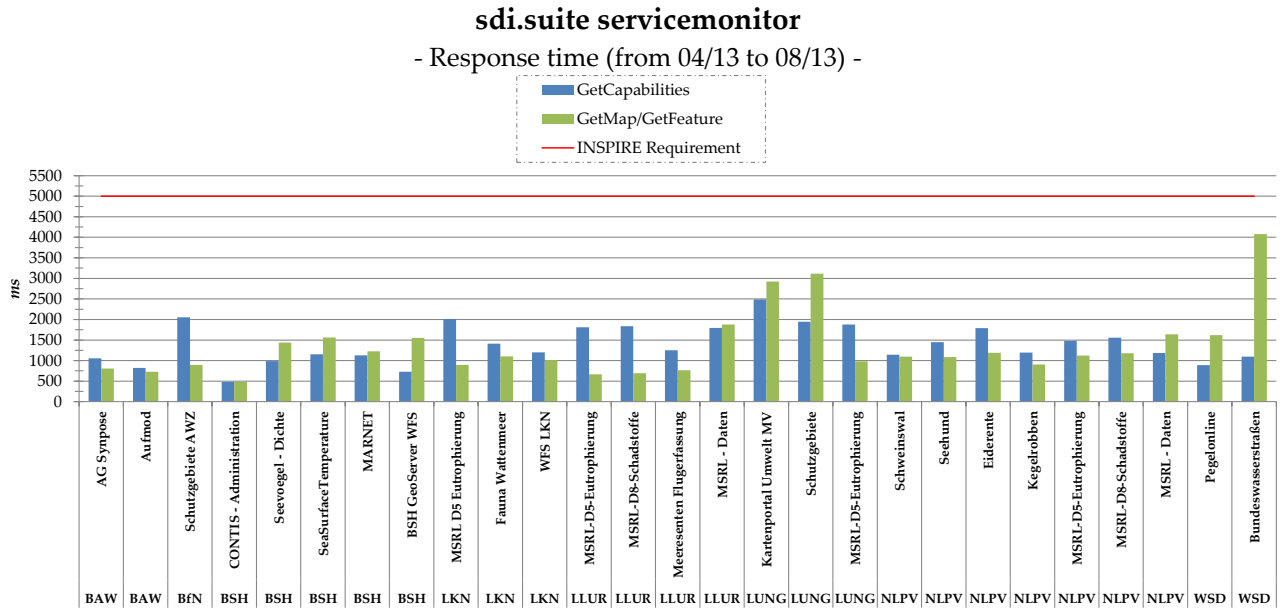


Figure 5.17: *Results for response time measured by sdi.suite serviceMonitor*

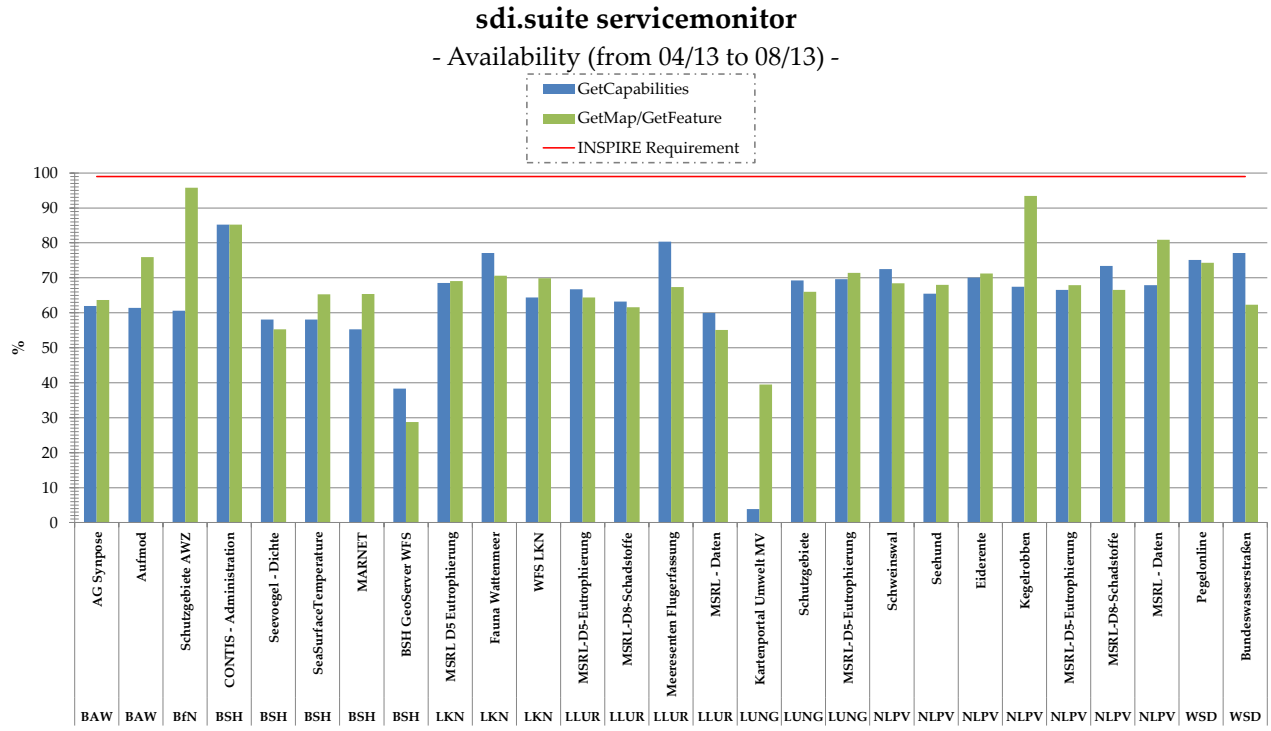


Figure 5.18: *Results for availability measured by sdi.suite serviceMonitor*

5.4.2.3 GDI-DE Testsuite

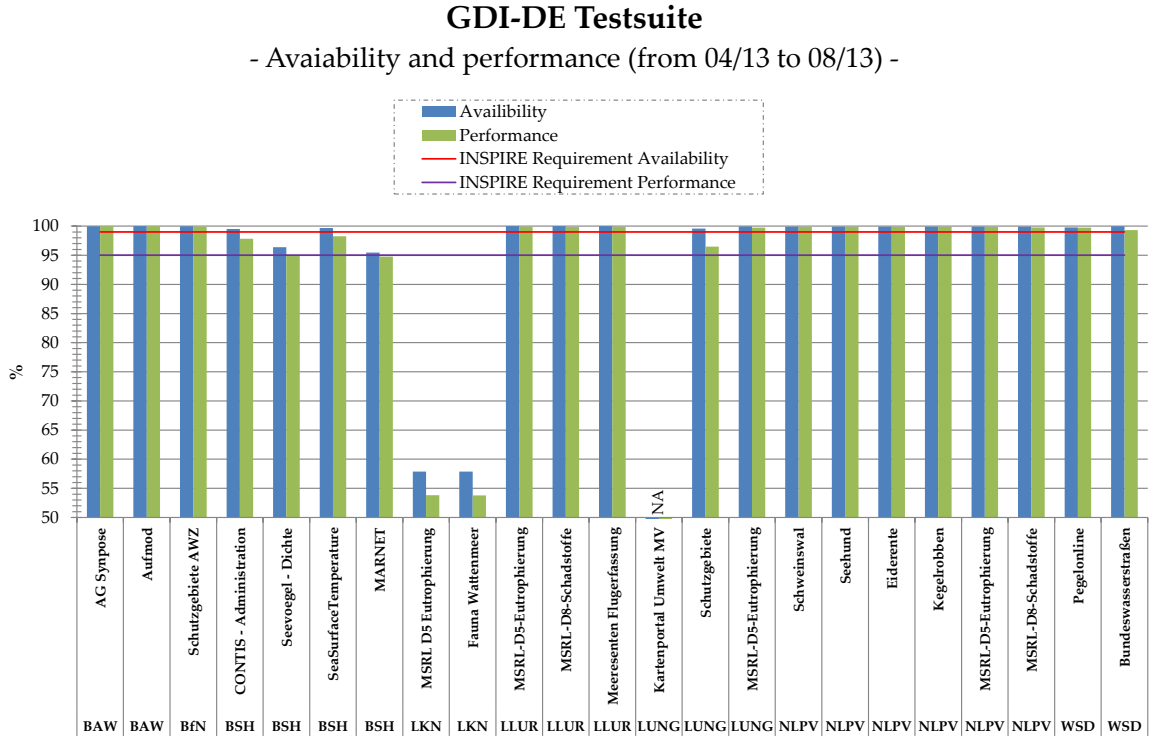


Figure 5.19: *Results for availability and performance by GDI-DE Testsuite*

The measurements with the GDI-DE Testsuite were made over a period of roughly four months (April to August 2013). In contrast to the sdi.suite serviceMonitor almost all the services were available at least 99 % of the time as depicted in figure 5.19. Only four services failed this INSPIRE requirement

(and one service could not be measured for unknown reasons). Although the INSPIRE requirement for performance is a response time of less than 5 000 ms to complete a GetMap request the Testsuite suggests that it also suffices that the response time is below 5 000 ms at least 95 % of the time. Only the two services from the LKN were not able to achieve this requirement.

5.4.2.4 Discussion and conclusions

As depicted in figure 5.20 according to the GDI-DE Testsuite almost all the services had an ***availability*** of at least 99 % with the exception of four. This is in stark contrast to the sdi.suite serviceMonitor where no service reached the INSPIRE requirement and the best service was available 85,24 % of the time which is way below the INSPIRE requirement of 99 %.

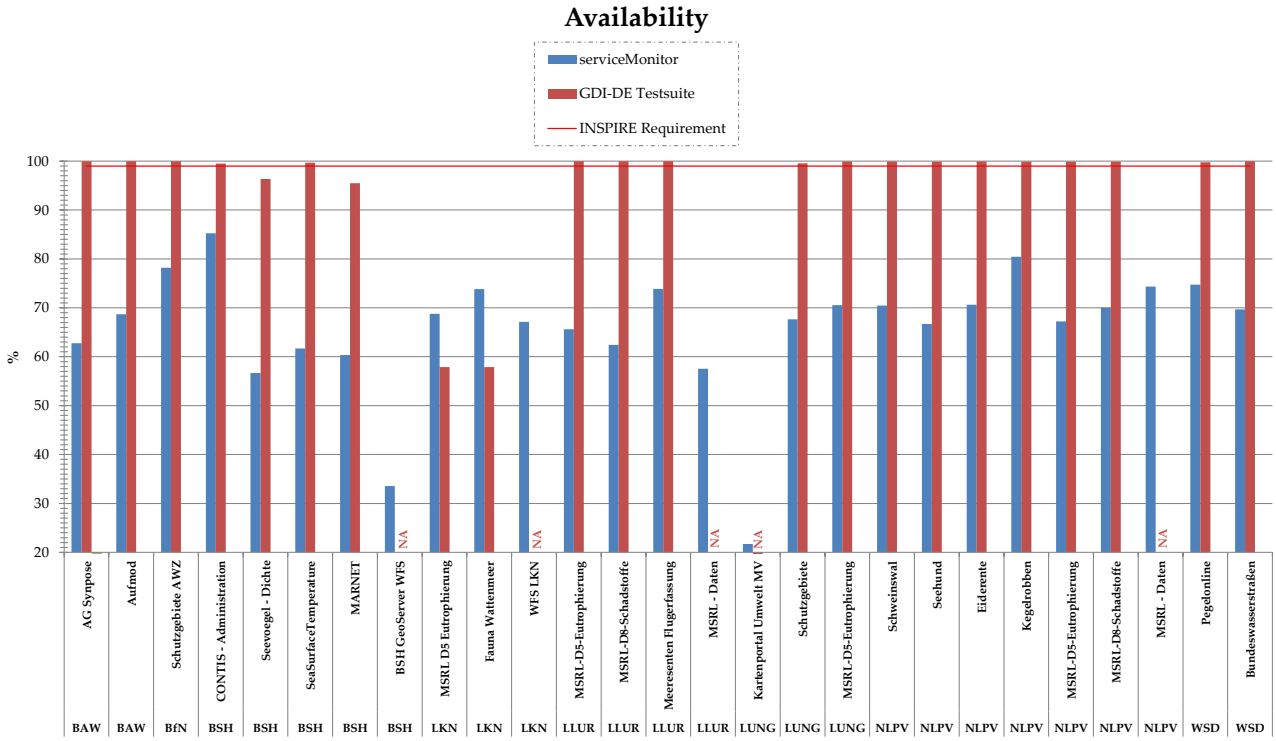


Figure 5.20: *Comparison of the monitoring results regarding availability*

Because of this contrast the Pearson product-moment correlation coefficient (PPMCC) is calculated to see to what extent the two samples are correlated. The PPMCC of r = 0,047 shows that the correlation between the two measurements is very weak (according to (Peter, 2006) r <= 0,2 means very low correlation and another source[15] even goes as far as to state that for r >= 0,01 and r <= 0,19 there is *" [. . .] no or negligible relationship"*). This is illustrated by the services "CONTIS - Administration" and "SeaSurfaceTemperature" by the BSH. While "CONTIS - Administration" has an availability of 85,24 % in the serviceMonitor it has 99,51 % in the Testsuite. "SeaSurfaceTemperature" has a comparable high value in the Testsuite of 99,66 % but only 61,7 % availability according to the serviceMonitor. With such a low correlation it is hard to trust the measurements. Either one of them is completely wrong or both are not to be trusted.

However, the measurements do not comply with INSPIRE requirements anyway because according to (INSPIRE Network Services Drafting Team, 2009, p. 14) the *"[. . .] response time is the time measured **on the server** [. . .] "*. The tools used for the measurements were running on a remote site. Another problem with the measurements is that it is unclear how many users were using the services and what they were doing at the time of measuring.

[15] faculty.quinnipiac.edu/libarts/polsci/statistics.html

A similar picture emerges regarding the ***response time***. Although the PPMCC is higher here (r = 0,064) compared to the availability it is still marginal. This is illustrated by the services "Schutzgebiete" and "MSRL-D5-Eutrophierung" both provided by LUNG. While both have a similar value of 688 respectively 590 ms when measured with SSC the serviceMonitor returned a value of 3 117 ms for "Schutzgebiete" and 980 ms for "MSRL-D5-Eutrophierung".

However, as figure 5.21 shows the measurements of the serviceMonitor and SSC agree that all the services are below the INSPIRE requirement of a response time of 5 000 ms.

Especially for the availability it would have made the results more comparable if exactly the same time periods would have been used. However, the main problem here is that the time periods were not equally long (roughly four months versus almost one month) due to the inability of SSC to return more than 1 000 records.

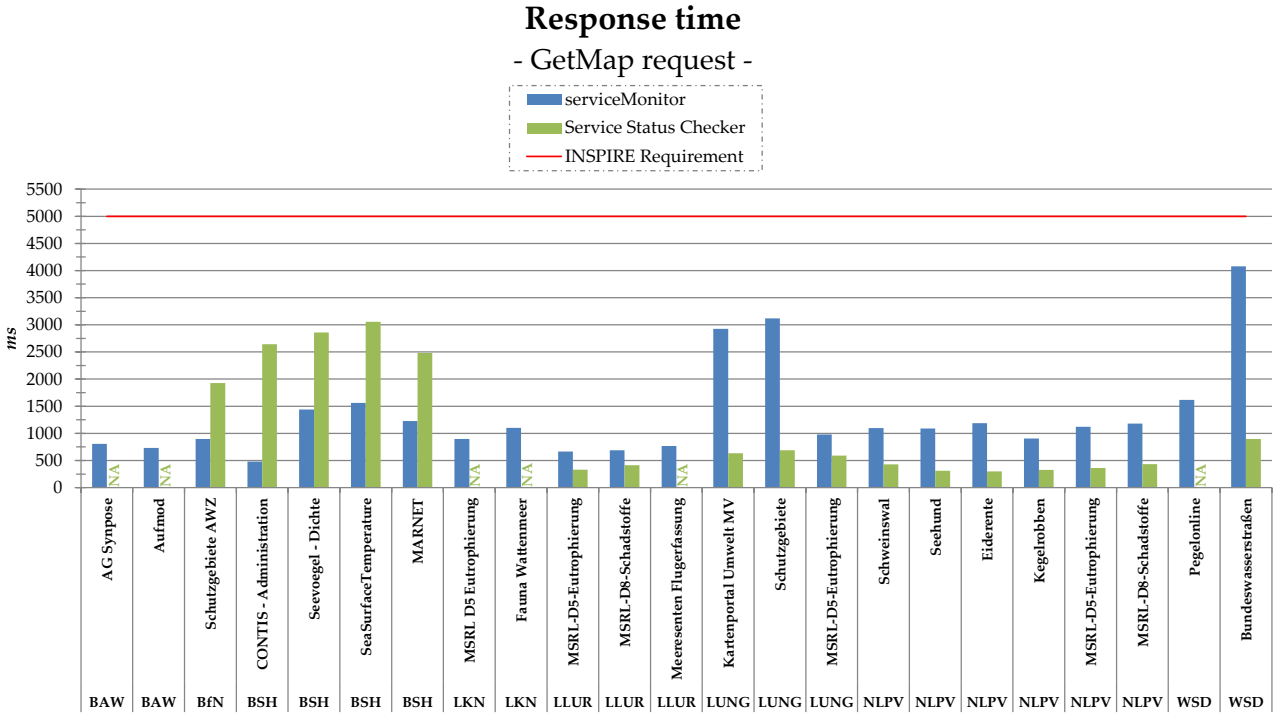


Figure 5.21: *Comparison of the monitoring results regarding response time*

Issues with measurements/monitoring in general

In addition to the availability issues already mentioned the variations due to the time of day affect the response time and the availability (for instance due to maintenance in the late evening or in the night). (Drerup, 2010, p. 18) arrived at the conclusion that these variations can make a difference of up to 4,25 %. (Drerup, 2010) also points out that the different request types have an impact on the response time but since this is fixed in INSPIRE regulations as GetMap respectively GetFeature this is not considered here. (Drerup, 2010) highlights an INSPIRE issue on page 27 by stating that *"[. . .] in the INSPIRE requirements the performance criteria for the Get Spatial Object operation is inaccurate. It can be fulfilled by requesting small datasets which does not meet practical applications."* This is not that important for the measurements in this thesis because there were very few WFS available and most tools were not able to handle WFS but is important to note for the future. (Cibulka, 2013) mentions a similar issue on page 31. He concludes that scale and location are linked to performance (response time). This means that requesting information with few data (e.g. points) results in lower response times. Two of the used tools (GDI-DE Testsuite and serviceMonitor) offer to formulate a request. The GDI-DE Testsuite is able to generate a GetMap request which was used for the QoS tests. The generated requests were also used for the serviceMonitor. Unfortunately SSC does not offer such

a possibility and generates a request on its own which cannot be altered. This might be a factor explaining why the results for the response time are so diverging.

It would be interesting to see whether there is a correlation between the results and the used server software. The used software is known for six hosts (encompassing 15 services) of which four use GeoServer, one ArcGIS Server and one MapServer. Unfortunately having only one sample each for a server does not constiture good scientific practice. Nevertheless table 5.3 shows the average means of availability and response time per server software. Of course availability has primarily to do with the host itself but the used server software can also crash itself. The pure response time has been primarily influenced by the internet connection of the hosts. That might not be true nowadays but what is more important is that the response time measured is the response time to a GetMap request which has to be processed by the server software which means that the used server software is the most important factor here. Table 5.3 shows that the response time results of GeoServer and MapServer are nearly the same which is backed up by (Müller & Mandery, 2009, p. 17) and (Aime & McKenna, 2009) who state that both deliver a similar performance. The table also shows that the response time of the ArcGIS Server host is significantly higher than the other two. (Nasser, 2009) states that *"[. . .] a lot of users might have complained about ESRI ArcGIS Server performance."*

Table 5.3: *Response time and availability of different server software*

	GeoServer	**MapServer**	**ArcGIS Server**
Availabilty [%]	77,3	85,2	89,1
Response Time [ms]	775	785	1 412

5.5 Visualization of Service Status Checker monitoring results

The preceding section (section 5.4) showed that results of SSC's service monitoring are only accessible in JSON format through its Application Programming Interface (API). Thus there is the need to visualize results of the Service Status Checker (SSC) to simplify evaluation of services with it.

5.5.1 SSC API and results

3.6.3.3 on page 72 pointed out that there are two main access methods for SSC: *Historical Tests* and *Live Tests*. Because the services should be monitored over a specific time span it only makes sense to make use of the *Historical Tests*. For these a user has to provide a list of services via ATOM feeds (or via a catalogue service). The services on this list are then batch tested twice a day. The results regarding performance and availability can be fetched through a web service (through an API using GET requests with output in JSON format) which is more detailed than the HTML service reports which is the second option to get the results.

The only two parameters needed to make a request to the API are `auth`[16] and `type`. Thus a simple request URL looks like this:

```
http://registry.fgdc.gov/statuschecker/api/v2/results?
auth=bb1022ba3a2821ca6369dd11b01a5d78&
type=wms
```

This only returns results from the last test run and not all test results that are available. To get more results and to specify a time span the optional parameters `from` and `to` are used. For a period of two weeks (e.g. `from=2013-08-01&to=2013-08-14`) 720 records are returned for the 24 services of the ATOM feed (twice a day times fifteen days times 24 services equals 720). The response is rather slow and hard to read because it is so long. But another parameter proves to be useful for that: `id`. This specifies a service and returns the records for this service only.

[16]Provided by SSC after registration.

The JSON output for each service is divided into a *summary* and a *tests* part after general information like the name, id and date of the test. The summary part provides a *currentSpeed* and *currentScore* while the tests part provides information about the HTTP server, getCapabilities and getMap request. The results for the test part are given through the parameter `success` which is either "1" (success) or "-99.99" (fail). Since these tests are not necessary to monitor performance and availability the focus lies on currentSpeed and currentScore.

According to its documentation[17] **currentSpeed** *"[. . .] is simply the time taken in seconds to do the test [. . .] "*. However, the source does not state what *the test* is. In a mail from July the 7th (2013) Michelle Anthony from the U.S. Geological Survey (USGS) stated that *"[. . .] only the GetMap test is defined as the "main" test [which] is the only test of the three that will be given a currentSpeed [. . .] "*. This means that currentSpeed represents the number of seconds a service takes to respond correctly to a getMap request. Furthermore the documentation (see above) states that the **currentScore** *"[. . .] is calculated as a factor of the Speed performance and the correctness of the response."* This factor is scored on a scale between 0 and 100. But how it is calculated explicitly remains unknown even after enquiry.

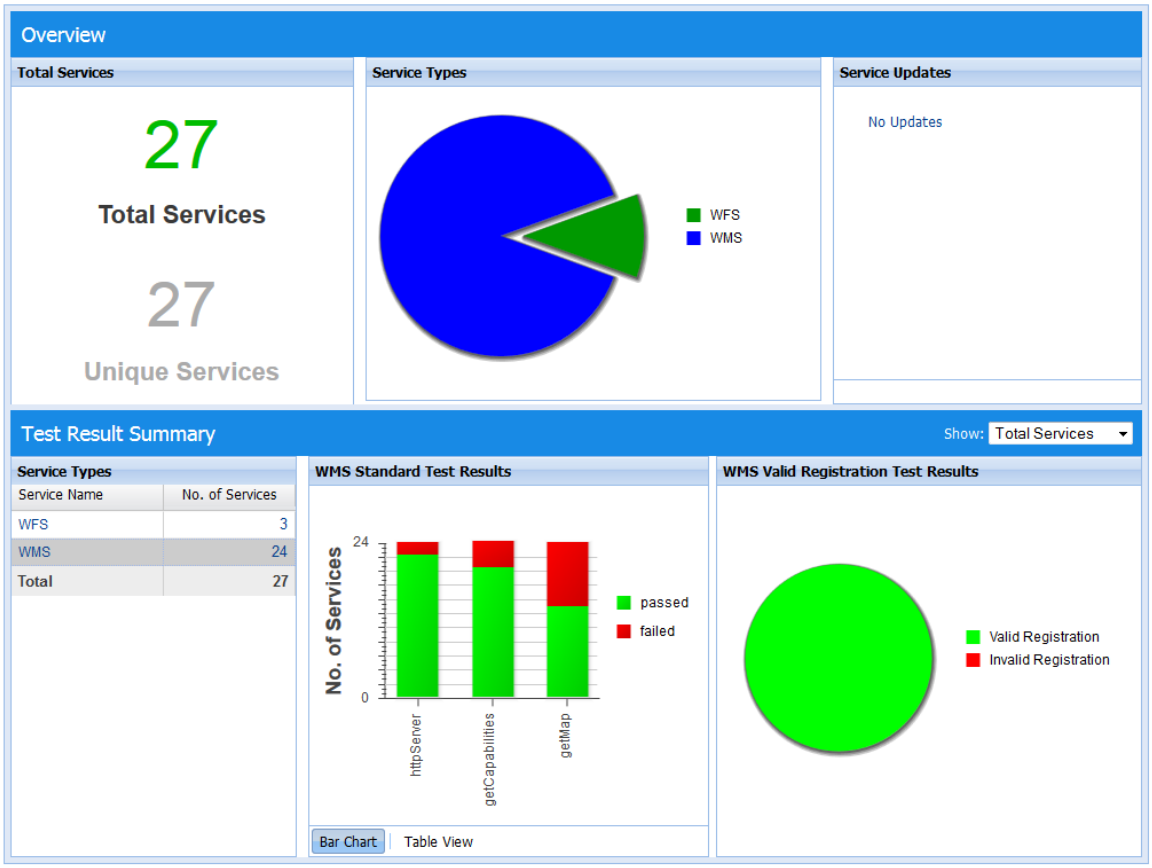


Figure 5.22: *SSC Dashboard showing 27 MDI-DE services*

Now that it is known how to query the API and interpret the results a visualisation is needed because the JSON output is confusing and complex for most users. The only visualisation found is the dashboard[18] depicted in figure 5.22. This only shows the number of services including the number of each service type (WMS and WFS) and a diagram or table showing how many services passed respectively failed which of the three tests. It shows neither the currentScore nor the currentSpeed which are the two most important parameters. Because such a feature is needed it was implemented for this thesis.

[17] http://registry.fgdc.gov/statuschecker/documentation.php#service-scoring
[18] http://registry.fgdc.gov/statuschecker/dashboard/index.php?feedID=mdides_32&serviceType=wms

To implement a visualisation for the two parameters currentScore and currentSpeed firstly a library respectively tool is needed to query the SSC API and get the results in JSON format. The method `getJSON()` from jQuery is very easy to use and accomplishes this task. jQuery also has a user interface library which offers a widget to pick a date and other widgets and functions. Since jQuery is a Javascript library in a second step a Javascript library is needed which can plot diagrams. jQuery offers a library capable of this called Flot. Lastly after the diagram is presented to the user he might wish to save the diagram. A function is needed to save the diagram directly as it is annoying to take screenshots, cropping them etc. This can be done with the HTML5 `<canvas>` element. Figure 5.23 shows the building blocks for the application called *SSCVisualizer*. The figure also shows the used libraries which are the foundation of *SSCVisualizer*[19].

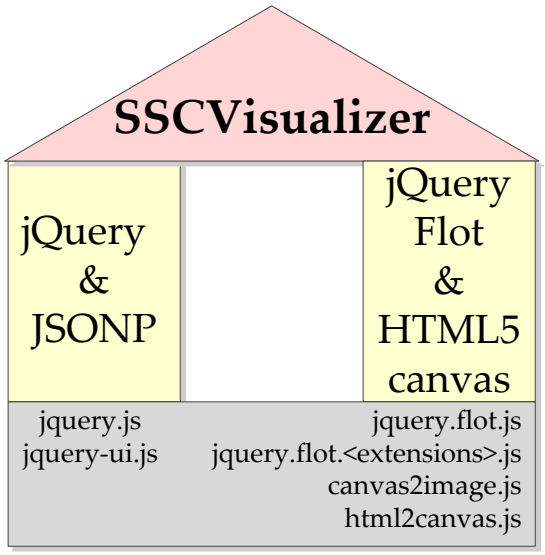


Figure 5.23: *Building blocks and foundations of the SSCVisualizer*

5.5.2 Same Origin Policy Problem

Figure 5.23 also shows that JSON with Padding (JSONP) is a building block, too. JSONP will be used to circumvent the *same origin policy* detailed in (Ihrig, 2012). While everything should work as depicted in figure 5.24 which means making an AJAX request using the `XMLHttpRequest` (XHR) object some place inside the HTML page this will not work because the XHR object is subject to the browser's same origin policy. This policy prevents access to resources from other domains than the origin (in this case the HTML page and its scripts are running on the domain `mdi-de.org` while the JSON data is requested from `registry.fgdc.gov`) to avoid security issues (the user could not know when a page is loading/requesting potentially harmful scripts from another domain).

However, the HTML `<script>` element is not subject to the same origin policy which means scripts from other domains can be loaded. JSONP uses this loophole and uses *dynamic script tag injection* depicted in figure 5.25. This technique creates a `<script>` element at runtime and after that its contents are loaded automatically. This means that the JSON response of `registry.fgdc.gov` is available as a script. Because JSON is a data format the client is not able to execute the script a callback function is needed. The JSON object is wrapped by the server in a call to such a callback function – padding it. Now the service's response is executable because it is a call to a function. In order not to confuse the name(s) of the callback function its name can be passed to the services with the parameter `callback`. This means the request URL will look like this now:

[19] Because of the libraries used and the extensive scripting the HTML page of *SSCVisualizer* is simple and short as listing A.9 on page 156 shows.

```
http://registry.fgdc.gov/statuschecker/api/v2/results?
callback=jsonpCallback&
auth=bb1022ba3a2821ca6369dd11b01a5d78&
type=wms
```

With such a request the service processes the two parameters `auth` and `type` like before, but returns a JSON object wrapped in a call to the function `jsonpCallback()` which provides the JSON data.

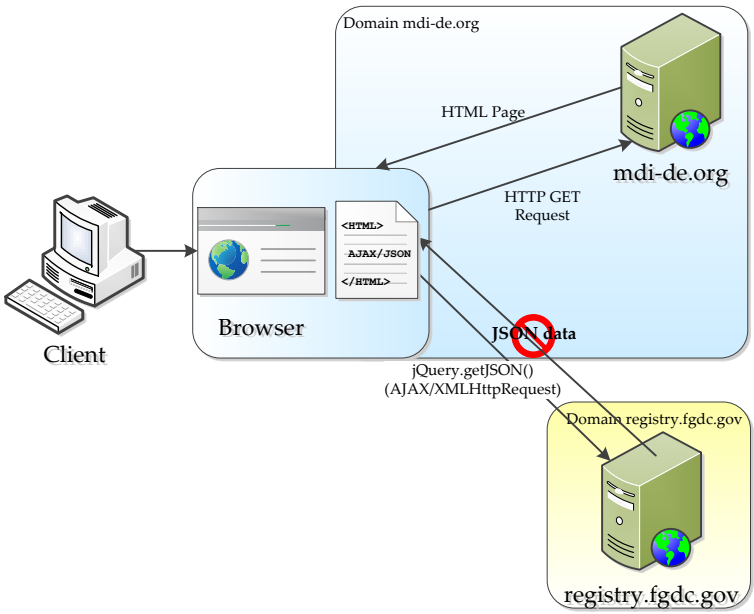


Figure 5.24: *AJAX (XMLHttp) request to get JSON data*

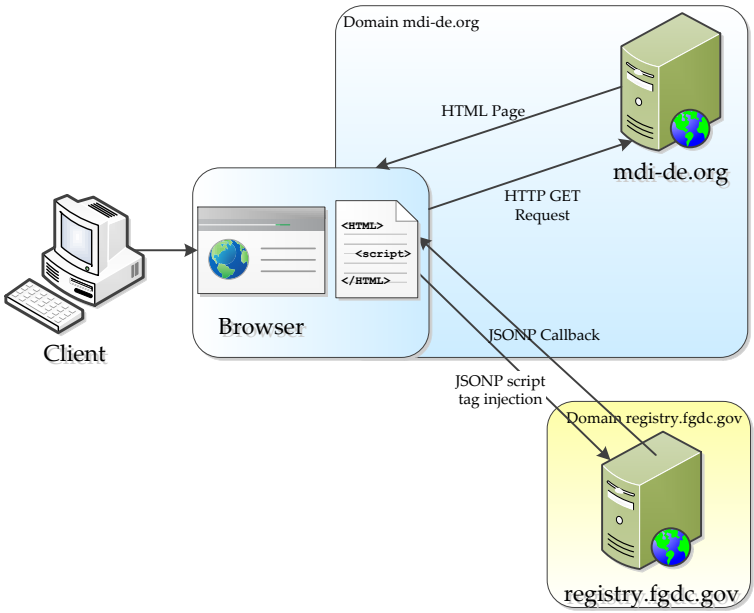


Figure 5.25: *JSONP request using dynamic script tag injection*

5.5.3 Creating Diagrams with Flot

In order to create a diagram *Flot* is used that – according to its website[20] – is *"[. . .] a pure JavaScript plotting library for jQuery, with a focus on simple usage, attractive looks and interactive features."* Its simplicity is demonstrated by how easy it is to create a diagram with Flot: only a HTML `<div>` element which holds the diagram and a call to the `plot()` function with the data is needed. The data and the function call is shown in listing 5.3 and the resulting simple diagram in figure 5.26. More complex data and feature requirements are handled by several plugins. There is a plugin for navigation (pan and zoom) that is useful when your diagram is getting big and hard to read when all data is displayed at once. There are also plugins to label the axes, handle time and categories.

```
1 $(function() {
2         //data array
3         var d1 = [[0, 3], [4, 8], [8, 5], [9, 13]];
4         //plot in div 'placeholder'
5         $.plot("#placeholder", [d1]);
6 });
```

Listing 5.3: *Basic code to create a diagram with Flot (based on http://www.flotcharts.org/flot/examples/basic-usage/index.html)*

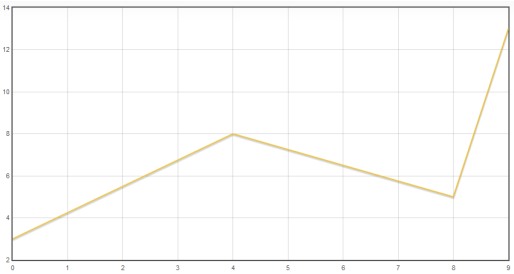


Figure 5.26: *Simple Flot diagram*

5.5.4 HTML5 <canvas>

According to the W3C[21] the HTML5 element `<canvas>` *"[. . .] provides scripts with a resolution-dependent bitmap canvas, which can be used for rendering graphs, game graphics, or other visual images on the fly."* The `<canvas>` element can be used to take "screenshots" of (parts of) web pages and save them to disk.

5.5.5 Implementation

Figure 5.27 depicts the implemented *SSCVisualizer* and shows that it is separated into three parts (A to C). Initially the user only sees section (1) of part A and both part B and C are empty. The user can input his authentication, the type of service and a time span. To supply a specific time span jQuery UI's `datepicker` was used to show a calendar and let the user pick the dates. Once the input fields are filled by the user he can click on *Fetch JSON & get Services*. This calls the function `jsonp()` which establishes a JSONP request and puts the JSON data into the callback function `jsonpCallback()`. This function puts the JSON data into a Javascript JSON object and calls `getServices()` at the end. This function iterates through the JSON objects and displays the available services as an option list (A (2)) from which a user can select the services he wants a diagram for.

[20] www.flotcharts.org
[21] http://www.w3.org/wiki/HTML/Elements/canvas

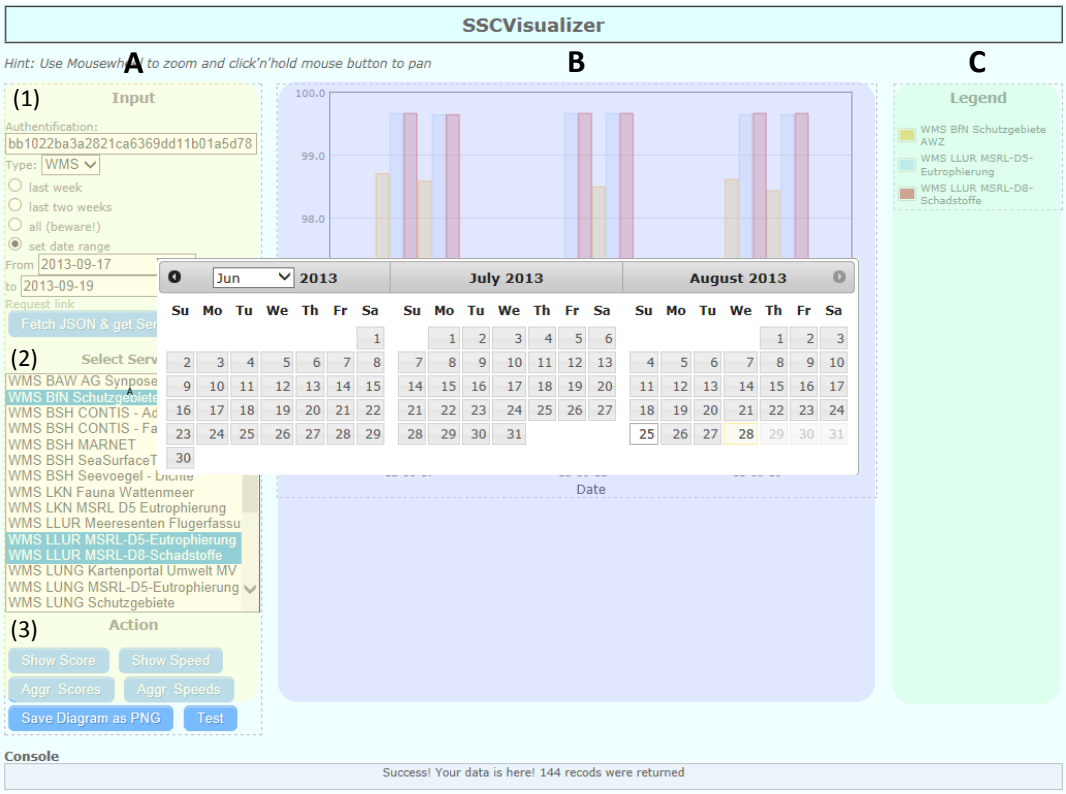


Figure 5.27: *SSCVisualizer (parts A to C highlighted using different colors)*

After selecting the services the user can choose if he wants a diagram for currentSpeed or currentScore by clicking on either one of the buttons in A (3) (*Show Score* respectively *Speed*). This shows all the recorded measurement results. The user can also choose *Aggr. Score* respectively *Speed* in A (3) to get a diagram with the mean values of the results over the time period specified.

The implementation of diagram creation for all the recorded measurement results is based on (Daws, 2012) after initially using Flot's categories plugin this inspired to use the time plugin instead. Firstly the function `showResults(type)` is called whereas `type` can be either "score" or "speed". This function immediately calls `getResults()`[22] which returns the dates, scores/speeds, id and names of the selected services as an array. After that `showResults(type)` begins to create a diagram with Flot, but the results have to be transformed first. The function `transformScores(results, type)`[23] converts the results array into a data array understood by Flot (different data format, width of the bars depending on the number of selected services has to be defined [function `calcBarWidth()`] etc.). Lastly `showResults(type)` makes sure every day is only displayed once (measurements are taken twice a day), configures the pan range (x axis, Flot's navigate plugin is used here), the minimum and maximum values of the y axis and labels both axes (Flot's axislabels plugin is used for this). The diagram is now displayed in part B and the accompanying legend in part C so it does not interfere with the bars of the diagram as figure 5.28 shows.

The implementation for diagrams with aggregated results is done in one function which immediately calls `getResults()` to get the records, too. It converts the results array into a data array understood by Flot and calculates the mean values in this step, too. Because the bars in the diagram (part B) are labelled a legend is not needed in part C which is depicted in figure 5.29.

[22] see listing A.6 on page 153
[23] see listing A.7 on page 154

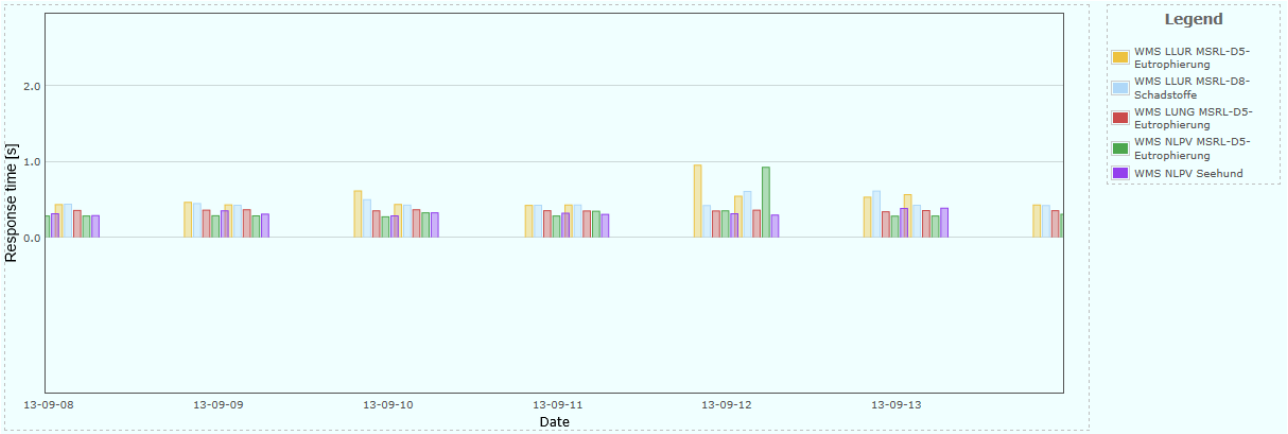


Figure 5.28: *SSCVisualizer displaying results (response time) for five services*

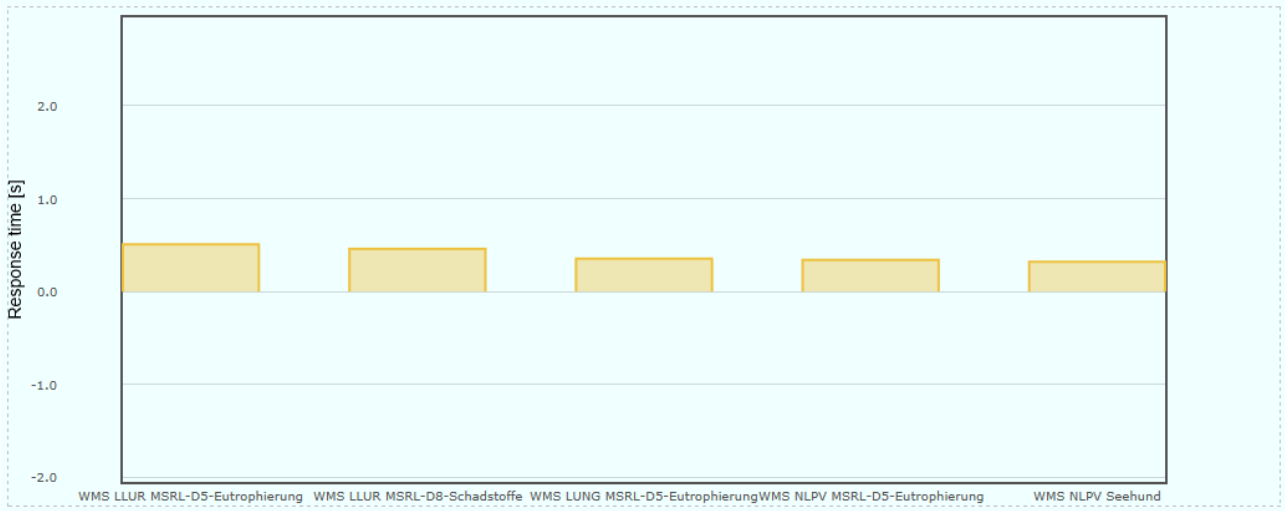


Figure 5.29: *SSCVisualizer displaying aggregated results (response time) for five services*

Based on the HTML5 `<canvas>` element it is very easy to save part of a website as image file (e.g. in PNG format) using the libraries *html2canvas*[24] and *Canvas2Image*[25] as listing 5.4 shows. *html2canvas* converts HTML elements (in this case a `<div>` with the id "scoreGraph") to a `<canvas>` element which is then saved as PNG using *Canvas2Image*. Since the diagram is useless or at least confusing without a legend both elements were merged and saved into one image by creating a new `<canvas>` element and using the `drawImage()` function of `<canvas>`[26].

```
1 function saveDiagram() {
2         html2canvas(document.getElementById("scoreGraph"), {
3                 onrendered: function(canvas) {
4                         Canvas2Image.saveAsPNG(canvas);
5                 }
6         });
7 }
```

Listing 5.4: *saveDiagram() based on HTML5's `<canvas>` element*

[24] html2canvas.hertzen.com
[25] nihilogic.dk/labs/canvas2image
[26] Listing A.8 on page 155 depicts the `saveDiagram()` method.

5.6 Requirements for a (marine) thesaurus

The prerequisite to setup services are data and according metadata sets. Especially for metadata annotation but also for services (e.g. keywords) and the MDI-DE portal itself (e.g. search function) thesauri are important. This is the reason why one aim of the MDI-DE project (see section 3.1) is the establishment of a marine thesaurus. A marine thesaurus is needed inter alia for users or experts to annotate their metadata with terms from a controlled vocabulary so that the usage of terms is consistent, typos are avoided and terms are linked to each other if they are related. This in turn improves discovery functionalities because when a user is looking for *shore* he will also get results mentioning *coastline*. Such a controlled vocabulary also makes it possible for the search functionality to provide an auto-complete function.

The existing word lists in the marine domain in Germany which were described in section 3.4 are in Excel format and in essence have columns for terms and definitions in different languages. On top of that some define relations to other terms with hints like “see” or “see also” or through hierarchies spanning several columns. To make use of the functionalities described above these lists have to be stored in a web-based thesaurus management tool of which some were presented in subsection 3.5.2. To select the one which is best suited for the needs of the to-be-built marine thesaurus subsection 5.6.2 first defines the requirements towards the tools and examines how they can fulfil these requirements afterwards.

In order to store the existing word lists in such a web-based thesaurus management tool they have to be converted into a format like SKOS. That is why subsection 3.5.1 introduced tools allowing conversion from text files to SKOS. Because the first step to build the marine thesaurus is the conversion of the existing word lists the tools are analyzed next (subsection 5.6.1).

5.6.1 Conversion Tools

The existing word lists in the marine domain which were discussed in section 3.4 are the foundation for the demands a conversion tools has to meet. For a conversion tool it is important that it:

(1) should not matter what the input text file looks like (generic) and
(2) supports hierarchies.

Almost all of the tools already fail the first requirement (1) because they expect some specifically formatted text files (Voc2Skos, Excel to SKOS/RDF conversion tool) or even a thesaurus respectively ontology format (Skosify, OWLtoSKOS, OBO and Zthes to SKOS Converter). The only tool which requires deeper inspection is OpenRefine. OpenRefine is good at converting arbitrary column-based text files into SKOS format but it would require specifically formatted text files to support hierarchies (2) which means that it is not possible to convert the hierarchies of the LHM (LANIS Habitat Mare) word list with it. On top of that there are some additional functions which would be very useful and nice to have in a conversion tool. These include:

- finding hierarchies automatically,
- finding related terms automatically and
- finding and setting links to matches in other thesauri.

Summing up it can be concluded that none of the existing tools meet the requirements listed above. However, they offer reasonable knowledge (e.g. the validation functions that Skosify offers) which will be used for implementing a tool (see section 5.7).

5.6.2 Web based Thesaurus Management Tools

To be able to fulfill the desired functionalities of the marine thesaurus which were mentioned in the introduction of this section (i.e. improving search and metadata annotation) the web based

thesaurus management tool used for the marine thesaurus has to have some key technical and thematic characteristics[27]:

- technical
  - expose terms respectively vocabularies through services
  - import thesauri in SKOS format
  - offer user management with roles and different rights
  - include collaborative content management
- thematic
  - support multiple languages (multilingual)
  - allow relationships between terms

Table 5.4 shows all these characteristics and if the tools from subsection 3.5.2 support them. It can be clearly seen that the features are almost the same. PoolParty is not Open Source (and thus not free) making PoolParty not an option. Although some features of MMI ORR are not known MMI ORR is not an option anyway. Firstly because you depend on them which means whenever they change something or if their services are unavailable there is nothing you could do about that. Secondly because we do not own the rights to all the existing vocabularies legal issues would be problematic. That leaves the two options TemaTres and iQvoc which are identical featurewise. An aspect in disfavour of TemaTres is their relative old-fashioned technology (mainly PHP) and their inadequate documentation (e.g. it is almost impossible to find information about their HTTP API). On the grounds of these considerations iQvoc is the tool which will be used to manage the vocabularies.

Table 5.4: *Comparison of web based thesaurus management tools*

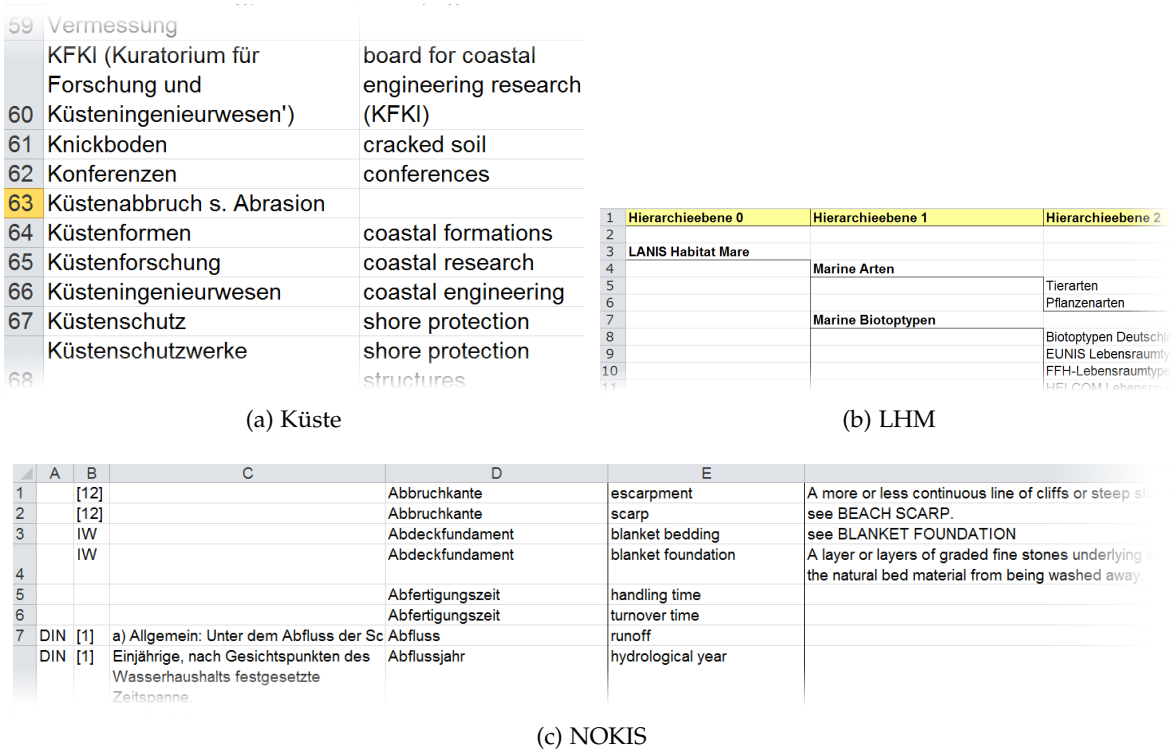
Tool / **Feat.**	*TemaTres*	*PoolParty*	*iQvoc*	*MMI ORR*
Services	HTTP API	HTTP API[28] and SPARQL	HTTP API	HTTP API and SPARQL
Multilingualism	+	+	+	+
Relationships between terms	+	+	+	+
Import	SKOS, tabbed txt	SKOS, CSV, Zthes[29]	SKOS N-Triples	CSV, Turtle, RDF/XML, N3, N-Triples
User management	+	+	+	NA
Collaborative content management	+	+[30]	+	NA
Open Source	+	-	+	NA

5.7 Converting vocabularies to SKOS for web usage

The preceding section (section 5.6) defined the requirements to build a marine thesaurus. These are the base for the actual implementation in this section. The existing vocabularies were introduced in section 3.4 on page 60 and are all in Excel format. Figure 5.30 depicts excerpts of the different vocabularies the application has to handle. Although these word lists are easy to comprehend for humans the lists cannot be used for indexing (metadata) or search because the Excel format does not allow computer systems to "understand" or make use of the vocabularies. However, converting word

[27] Being Open Source is an additional characteristic due to lacking funds for commercial software.
[28] http://poolparty.biz/de/skos-without-sparql-poolparty-skos-api/
[29] https://grips.semantic-web.at/pages/viewpage.action?pageId=40437853
[30] http://poolparty.biz/poolparty-functionalities-features-at-a-glance/

lists into SKOS format which was introduced in 2.4.2.2 on page 37 changes that. Vocabularies in SKOS form can be handed over to SKOS management tools like TopBraid Enterprise Vocabulary Net or iQvoc. Through the use of such tools the vocabularies can be visualized or maintained through a web browser which enables specialists to contribute their knowledge to a vocabulary.



59	Vermessung	
	KFKI (Kuratorium für	board for coastal
	Forschung und	engineering research
60	Küsteningenieurwesen')	(KFKI)
61	Knickboden	cracked soil
62	Konferenzen	conferences
63	Küstenabbruch s. Abrasion	
64	Küstenformen	coastal formations
65	Küstenforschung	coastal research
66	Küsteningenieurwesen	coastal engineering
67	Küstenschutz	shore protection
	Küstenschutzwerke	shore protection
68		structures

(a) Küste

	Hierarchieebene 0	Hierarchieebene 1	Hierarchieebene 2
1			
2			
3	LANIS Habitat Mare		
4		Marine Arten	
5			Tierarten
6			Pflanzenarten
7		Marine Biotoptypen	
8			Biotoptypen Deutschl
9			EUNIS Lebensraumty
10			FFH-Lebensraumtype

(b) LHM

	A	B	C	D	E	
1		[12]		Abbruchkante	escarpment	A more or less continuous line of cliffs or steep sl
2		[12]		Abbruchkante	scarp	see BEACH SCARP.
3		IW		Abdeckfundament	blanket bedding	see BLANKET FOUNDATION
4		IW		Abdeckfundament	blanket foundation	A layer or layers of graded fine stones underlying the natural bed material from being washed away
5				Abfertigungszeit	handling time	
6				Abfertigungszeit	turnover time	
7	DIN	[1]	a) Allgemein: Unter dem Abfluss der Sc	Abfluss	runoff	
	DIN	[1]	Einjährige, nach Gesichtspunkten des Wasserhaushalts festgesetzte Zeitspanne.	Abflussjahr	hydrological year	

(c) NOKIS

Figure 5.30: *Examples for existing German marine vocabularies (excerpts)*

Subsection 5.7.1 will lay the foundations for the implementation, looks at alternatives and what can be learned from existing tools. Afterwards subsection 5.7.2 will develop a concept for the implementation which will be described in subsection 5.7.3.

5.7.1 Foundations

Subsection 3.5.1 on page 62 introduced tools that can be used to convert a given format into SKOS format. Subsection 5.6.1 on page 134 concluded that none of the existing tools fulfil the requirements proposed in section 5.6 on page 134. However, another way to conduct such a conversion without using an existing tool is using Extensible Stylesheet Language Transformations (XSLT). Through the use of XSLT code and an XSLT processor the desired results can be achieved. However, this approach is not generic enough because the XSLT code would have to be (re-)written for every word list which means reusability would be low if there is more than one vocabulary. Furthermore additional functionalities (see subsection 5.7.3) would not be possible using XSLT.

Although no tool was able to fulfil the requirements there may be lessons that can be learned from these existing implementations. *Skosify* is not suitable for the conversion of arbitrary word lists (e.g. in Excel or CSV format) into SKOS because it accepts files only in formats of the semantic web (RDF and OWL) and such conversions could easily be done with frameworks like Sesame 2 as well. However, it provided some usable knowledge through its processing steps. Steps such as making sure the vocabulary has a `skos:ConceptScheme` and performing validations such as making sure that there is only one `skos:prefLabel` per language are important for an own implementation.

Open Refine was the tool closest to fulfilling the requirements but cannot handle hierarchies or additional functionalities (see subsection 5.7.3). However, letting the user select the resources and

literals and what they will become in the RDF document can be considered for an own implementation as well as defining an URI. *Zthes* also lets the user specify a URI. Users can specify a base name space and a concept scheme ID which prepends the name space in the URI which is interesting, too. In Voc2skos the URI can be defined using an `ontologyURI` element in the preamble of a CSV (Comma-separated values) document. Furthermore *Voc2skos* even supports hierarchies through the use of indentation in a CSV file. However, the problem is that the user has to alter the data (change headings, indentation etc.). The same problem can be found in the *Excel to SKOS/RDF conversion tool* which requires the data to adapt and not the other way round. Both approaches violate the "rule" of data life time (Christl, 2013) *"Your software will go away. Your data is going to stay."*

In addition to the points made in subsection 5.6.2 on page 134 *iQvoc* is used as thesaurus web management tool because one of the project partners of the MDI-DE project is the Federal Environmental Agency (UBA) and they initiated the development of iQvoc. Furthermore iQvoc – just like SKOS itself – builds upon the four principles of the Linked (Open) Data Concept by (Berners-Lee, 2003):

- Use URIs as names for things.
- Use HTTP URIs so that people can look up those names.
- When someone looks up a URI, provide useful information, using the standards (RDF*, SPARQL).
- Include links to other URIs, so that they can discover more things.

5.7.2 Concept

The preceding section underlined what is important when setting up or converting a SKOS document:

(1) Do not change the data – the tool has to adapt to the data
(2) Let the user specify what SKOS properties the resources and literals, i.e. column headings, will become
(3) Setup the Document and make sure it uses a `skos:ConceptScheme` and HTTP URIs

However, the first step for a tool that converts vocabularies to SKOS format of course is to import and load the file. The next step is to setup the SKOS document which means to specify name space locations for SKOS and RDF in the preamble of the document. Subsequently a `skos:ConceptScheme` (note the prefix `skos` that is now usable) will be specified that uses an HTTP URI to which all the concepts will point. In order to do this the user has to specify a base URI which typically is the URL of the server that is hosting the thesaurus, e.g. http://www.example-thesaurus.com. Now all the concepts could point to http://www.example-thesaurus.com#ExampleConceptScheme and are available as http://www.example-thesaurus.com#ExampleConcept. The problem with URIs like that and the combination of more than one vocabulary is that for example the term "beach" might appear in more than one vocabulary and it cannot be stored as http://www.example-thesaurus.com#Beach multiple times. This means that an additional term is needed to narrow down the vocabulary the term originates from. Now the term is available under http://www.example-thesaurus.com/#Beach_LHM for example.

After the import of the vocabulary and the setup of the SKOS document the user is presented with the column headings and selects the suiting SKOS properties and their language for each column. Additionally the user may select the hierarchy level of a column for vocabularies like LHM depicted in figure 5.30b on page 136.

Afterwards the vocabularies are read line by line which fills the SKOS document with concepts (each line is a concept). Relations between concepts of the same vocabulary and/or other vocabularies can be incorporated after or within the preceding step. The final step is the saving of the document to disk in a semantic format like RDF XML, Turtle or N-Triples whereat iQvoc only accepts imports in N-Triples format.

5.7.3 Implementation (JSKOSify)

Based on the concept from the preceding section six steps evolved which need to be implemented:

1) Import vocabulary
2) Setup Document
3) User specifies what SKOS properties the columns will be
4) Fill document
5) Relationships (Hierarchies, Matches in other thesauri etc.)
6) Save document

These steps are reflected in the overview of the implementation depicted in figure 5.31 which also shows the division into two main classes. One class (`JSKOSifyImpl`) implements the steps 1, 2 and 4 to 6 and the other class (`JSKOSifyGUI`) uses the functions of `JSKOSifyImpl` and provides the user a Graphical user interface (GUI) to carry out step 3 [31]. The three green helper classes will implement parts of step 5. The next sections will detail every step and the whole code is available on GitHub[32].

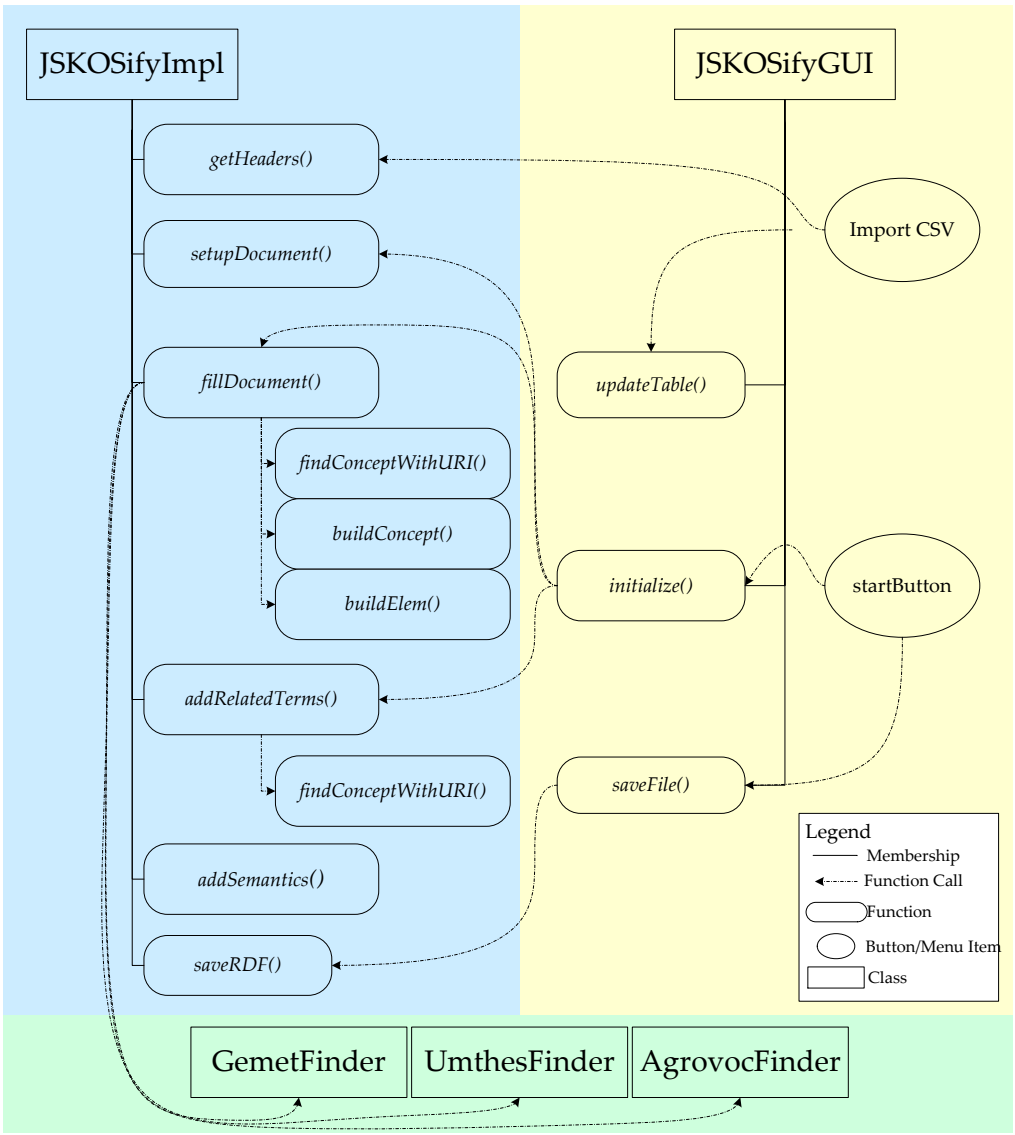


Figure 5.31: *JSKOSify overview (classes and functions)*

[31] see listing A.13 on page 159 for exemplary GUI class methods
[32] https://github.com/Sicky/JSKOSify

1) Importing vocabularies

Although the existing vocabularies are in Excel format they were converted to CSV (Comma Separated Values) format because no Excel specific functionalities were used. Furthermore Excel is a proprietary format and there are free libraries which can work with CSV. The free and open source Java library CSVReader[33] was chosen that imports CSV files row by row based on the column headings which suits the implementation approach as all word lists have column headings because otherwise the lists would be hard to interpret.

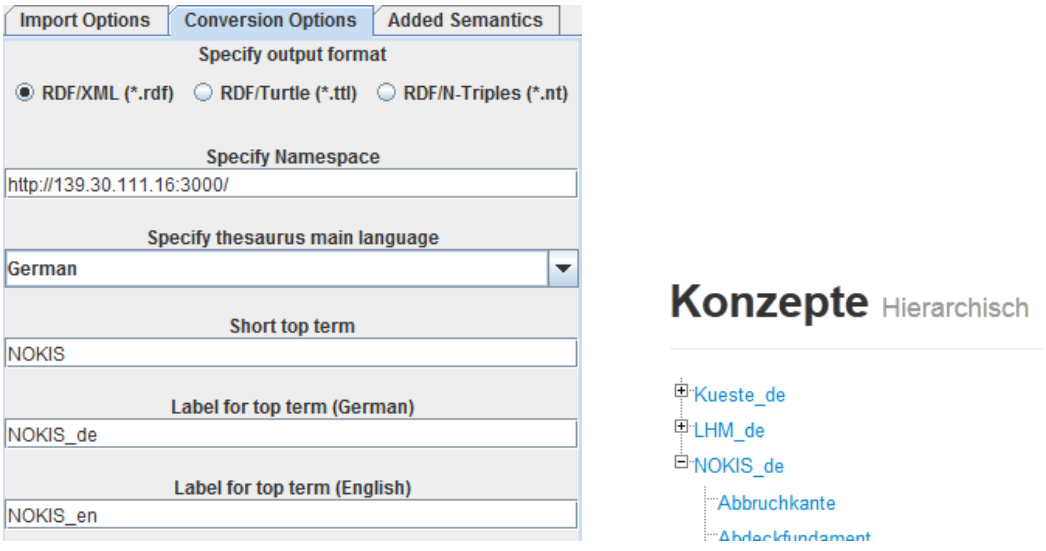
The GUI of JSKOSify calls a function called *getHeaders()* when importing a CSV file. This function uses CSVReader for the table presented to the user (to be able to assign the SKOS properties to the columns, see figure 5.33 on page 140) that is constructed by the function *updateTable()*. CSVReader will also be used in the function *fillDocument()* that will be detailed later.

2) Setting up the SKOS document

The function *setupDocument()* is actually called after the user assigned the SKOS properties but logically it comes first. Firstly it creates a JDOM document with namespace declarations for RDF and SKOS (lines 3 and 4 in listing 5.5). Secondly and lastly it creates a superordinate concept (lines 6 to 10 in listing 5.5) to act as the root in the hierarchy of the word lists (shown in figure 5.32b). This is created based on user input which is depicted in figure 5.32a.

```
1  <?xml version="1.0" encoding="UTF-8"?>
2  <rdf:RDF
3    xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
4    xmlns:skos="http://www.w3.org/2004/02/skos/core#">
5
6   <rdf:Description rdf:about="http://139.30.111.16:3000/#NOKIS">
7    <skos:prefLabel xml:lang="de">NOKIS_de</skos:prefLabel>
8    <skos:prefLabel xml:lang="en">NOKIS_en</skos:prefLabel>
9    <rdf:type rdf:resource="http://www.w3.org/2004/02/skos/core#Concept"/>
10  </rdf:Description>
11 </rdf:RDF>
```

Listing 5.5: *Document after initial setup*



(a) GUI to specify settings

(b) Root concepts in iQvoc

Figure 5.32: *View of JSKOSify GUI and iQvoc (both excerpts)*

[33] http://www.csvreader.com/

The user specifies a namespace (or the base URI) and a short term that describes the vocabulary. Following listing 5.5 (concept e.g. available as http://139.30.111.16:3000/#Beach_NOKIS) the namespace is `http://139.30.111.16:3000` and the short term is `NOKIS`. That way the user specifies the URI for his concepts in this step. The superordinate concept is only built because neither a `skos:ConceptScheme` nor a `skos:Collection` work in the iQvoc version used.

3) Assigning SKOS properties

Using CSVReader the function *updateTable()* constructs the table depicted in figure 5.33 after the import of the CSV file. The column headings of the CSV file are placed in the first column of the table. In the next column the user can select the SKOS property to the column of the CSV file. For example column *definition_en* will probably contain an English definition of a term. That is why the user selects `skos:definition` in the second column and "English" in the next (third) column. For vocabularies like LHM (depicted in figure 5.30b on page 136) it is necessary that the user specifies the hierarchy level of the column (of the CSV file) which is done in the fourth column. A hierarchy level of 1 for example means that his column does not have any broader concepts but (most certainly) several narrower concepts. Lastly the user selects if the specific column should be used for the URI. This means if the column with an English term (in the example *begriff_en*) should be used for URI construction the concept is available under http://139.30.111.16:3000/#Beach_NOKIS for example. If the column with a German term should be used then the URI would be http://139.30.111.16:3000/#Strand_NOKIS in this case.

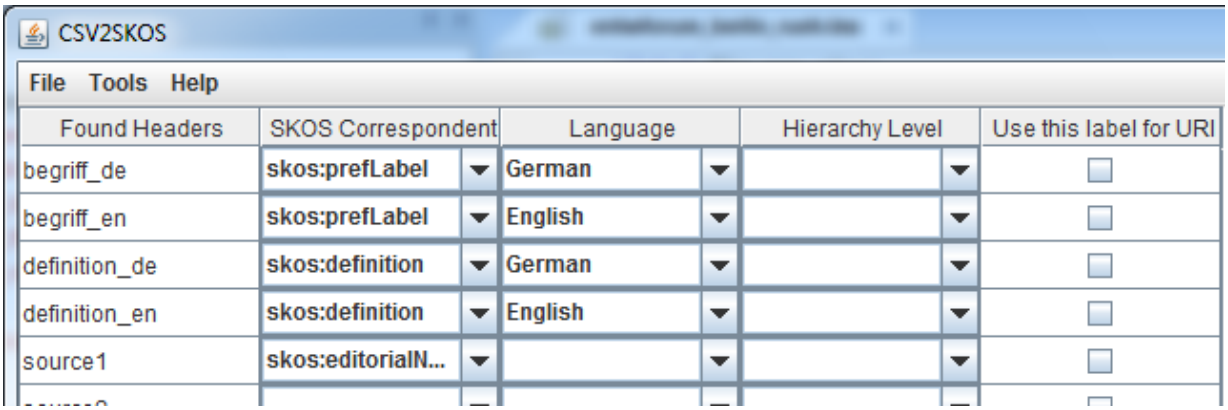


Figure 5.33: *GUI to assign SKOS properties*

4) Filling the document with concepts

After the user assigned the SKOS properties to the columns of the CSV file the **Start** button can be clicked. This will setup the document firstly (see step 2) and then fills the document with concepts using the *fillDocument()* function that iterates through the CSV file line by line. Because some word lists have multiple rows for the same term (concept) it has to be checked whether a concept already exists or not using the function *findConceptWithURI()* (see line 4 in listing 5.6). If it already exists the existing concept will be used otherwise a new concept will be created using the function *buildConcept()*. When the concept is created a `skos:broader` element is added which points to the superordinate concept (for correct representation in tree view). After that an element will be created for each column (of the CSV file, see line 10 to 15 in listing 5.6) using the function *buildElem()*[34] and added to the concept. Continuing the example (the column with an English term) the resulting element could be:

```
<skos:prefLabel xml:lang="en">beach</skos:prefLabel>
```

Instead of a `skos:prefLabel` it also could be a `skos:altLabel` depending on what the user had selected. The last step adds the concept to the root of the document (see line 17 in listing 5.6).

[34] see listing A.10 on page 157

```
while (csvReader.readRecord()) {
 //Check if there is an existing concept with that URI
 Element oldElem = findConceptWithURI(doc, URI4NewConcept);
 //If there is use this concept
 if (oldElem!=null) concept = oldElem;
 //Otherwise create a new concept
 else concept = buildConcept(topTerm, namespace, mainLangShort);
 //iterate through the columns of the CSV file
 for (int i=0; i<columns; i++){
  //Build a new element for the concept
  Element element = buildElem(elemTerm, skosCorrespondent, language);
  //Add new element to the concept
  concept.addContent(element);
 }
 //Add concept to the document
 doc.getRootElement().addContent(concept);
}
```

Listing 5.6: *Reading the CSV file line by line and constructing new concepts with the information (mockup)*

5) Adding relationships

There are three main tasks and some detailed tasks concerning relationships that can be carried out if the user selects them (see figure 5.34):

1) Matches in other thesauri (GEMET[35], UMTHES[36], AGROVOC[37])
2) Hierarchies within a vocabulary
3) Relationships within a vocabulary

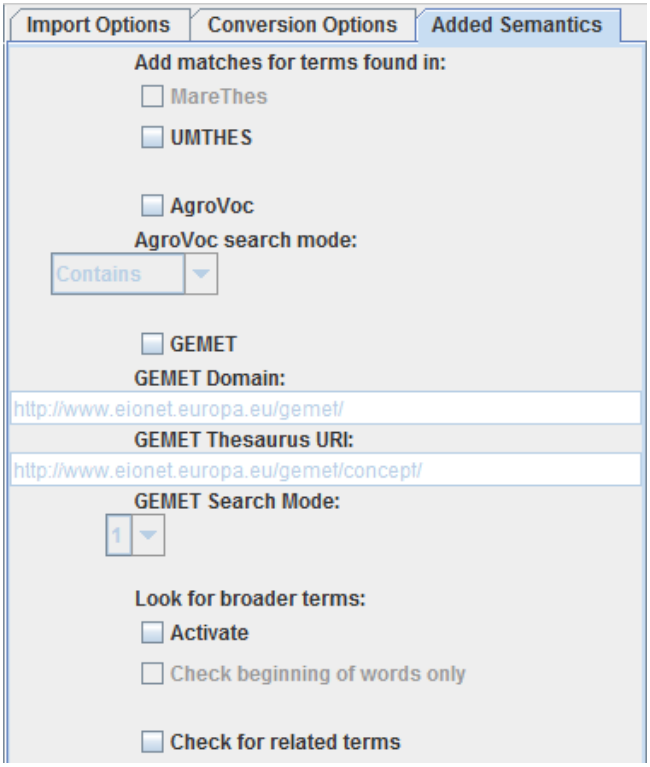


Figure 5.34: *GUI to select relationship taks*

[35]http://www.eionet.europa.eu/gemet/
[36]http://data.uba.de/umt/
[37]http://aims.fao.org/standards/agrovoc/

Task 1) is carried out using three helper classes[38] which send a HTTP GET query using the API of one of the thesauri. If a match is found an element is created which looks like this:
`<skos:exactMatch rdf:resource="http://[..]/concept/2969"/>`.
Afterwards this element is added to the concept.

Task 2) looks for terms in the document respectively vocabulary which are obviously narrower terms (e.g. "beach scarp" has "beach" as broader concept). The function *addRelations()*[39] iterates through the whole document and checks a part of the term against all the terms in the document. If a match is found a `skos:broader` element is added to the narrower concept and the existing `skos:broader` element (pointing to the superordinate concept) is deleted.

Task 3) works similar and was implemented for the vocabulary "Küste" because it expressed relationships between terms with the declarations "s." (see) and "s.a." (see also). The function *addRelatedTerms()* checks for every row if there is a "s." or "s.a." and if there is, it uses the function *findConceptWithURI()* to see if the concept respectively term pointed to exists in the document. If the concepts pointed to exists a `skos:related` element is created and added to the concept.

6) Saving the document

The last step of the conversion is to save the document to disk. After finishing the steps before the GUI asks with a JFileChooser dialog were to save the file and calls the function *saveRDF()*[40] that saves it in the format the user specified (see figure 5.32a). There is also an additional function called *convertRDF2NT()* which can be called from the menu bar. This converts a given RDF XML file to RDF N-Triples format which is the only format that iQvoc accepts.

[38] see listing A.14 on page 160 for one of these
[39] see listing A.11 on page 157
[40] see listing A.12 on page 158

6 Future prospects

The conceptual work and the selected implementation aspects introduced in this thesis on the one hand are approaches only and on the other hand are in prototypical stages. This leaves plenty of room for expansions, improvements and adaptations of these aspects.

Evaluation of existing MSDIs Although the framework is quite advanced some adaptions could be needed in the future. It may emerge that more indicators are needed because there are still things that were not thought of. It could also be the case that one or more indicators prove to be unnecessary, ambiguous or too hard to measure. That way indicators may be left out or be defined differently.

However, such things will only show when the framework is applied to MSDIs. Ideally this will be done by experts in the marine field or maybe even operators or users of a certain MSDI. In order to do this a web site with an online form could be used. The web site could also aggregate the results and compute an average. For such a feature it is important to inform experts and announce it broadly.

Reference Model After the project was finished works on the reference model have also ended which leaves no room for future changes or improvements. However, it appeared that because of the massive coordination efforts more tools for collaborative coordination and working on documents would simplify the set-up of a reference model. Furthermore interactive tools for collaborative data modelling would improve such processes for instance.

Marine thesaurus The tool *JSKOSify* proved to be useable to convert vocabularies to SKOS format. However, there are several additional features imaginable such as editable table entries, for this feature an autocomplete functionality, save and restore profiles and using the REST API of iQvoc directly from JSKOSify to import concepts or vocabularies into iQvoc. Furthermore existing features could be improved such as the automatic lookup of broader terms which could be augmented with parameters the user can change. The lookup of relationships in other vocabularies than GEMET and AgroVoc could also be augmented with parameters. Furthermore error handling could be improved here when other vocabularies return unusable responses or are not available.

After *JSKOSify* was used to convert the existing vocabularies into iQvoc experts in the marine domain have to coordinate the harmonization of them to form a single marine thesaurus for Germany. Furthermore the presentation of thesauri could be improved in the futures. An example for a more interactive thesaurus is www.visualthesaurus.com which is depicted in figure 6.1.

Evaluation of MDI-DE services The evaluation of the services showed that the different tools are not comparable. However, it would still be interesting to see if their results change over time (i.e in the future) especially when the systems have left their prototype stage and productive. Furthermore it would be interesting to know if changes on the server side can influence the monitoring results. But this would be possible only if one has access to the server(s) itself.

The web tool *SSCVisualizer* is quite advanced but the diagram presentation of the results may could be improved, especially for aggregated results. Furthermore the activation and deactivation of the buttons is still troublesome. As it is a web tool testing it with different web browsers is also needed to make sure it works with other browsers as well.

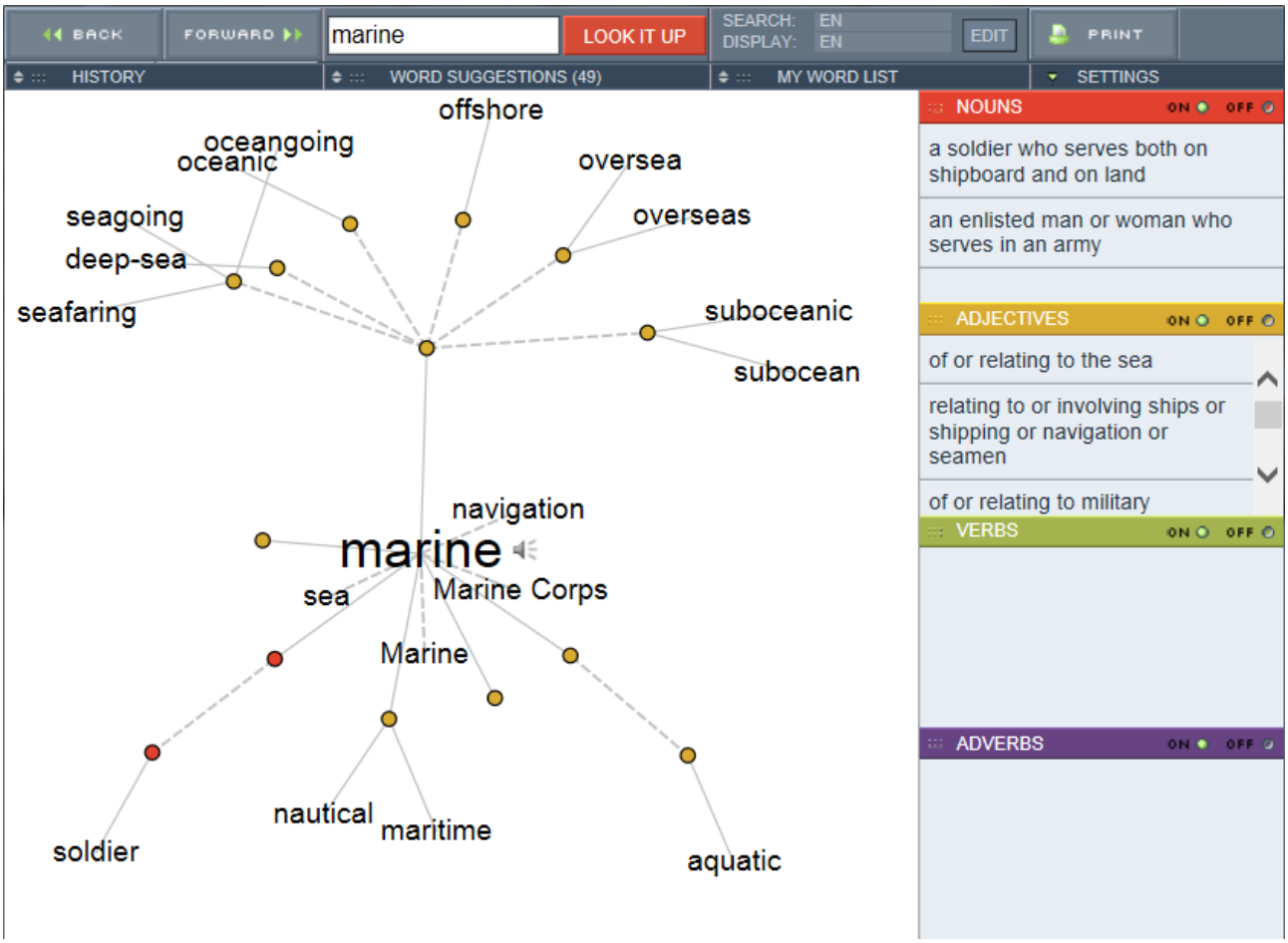


Figure 6.1: *Lookup of the term marine on visualthesaurus.com*

Further prospects The IT sector and thus the web are always subject to change and ongoing improvements. Therefore it will be interesting how MSDIs will develop. The United States of America for instance show a lot of effort integrating marine (spatial) data into their national SDI which itself was integrated into their general public data portal *data.gov*. This opens up the question if other countries will follow as this strategy simplifies data access because it is like a one-stop for all available data.

A similar route could be taken by data and metadata which means that both may also grow together. The increasing usage of WFS nowadays is a hint of this way because they offer the metadata attached to the data much more directly than a WMS. That way the user does not need extensive search for metadata. Thus WFS will (hopefully) be used more and in conjunction with WMS in the future. Furthermore WFS might be accompanied by more and more gazetteer and thesaurus services that will ease data lookup and provide clarification of terms. In addition services like the Web Processing Service (WPS) and Sensor Observation Service (SOS) amongst others will become more relevant.

At the present time Open Data gains significance. With services such as WFS and SOS raw data can be published and because of the Open Data movement this will be used more in the future. On the one hand this may relieve authorities because they do not have to modify (aggregate etc.) the data anymore. On the other hand this enables companies and citizens to create new things out of the data which may contribute to the society's progress. This could result in more mobile apps with specialized thematic content. However, it remains to be seen whether data acquiring institutions make use of Open Data or are allowed to make use of Open Data or not. Furthermore quality assurance concerns may speak against the use of Open Data.

Appendices

A Selected listings

A.1 Analysis of existing data sets

A.1.1 Creation of a database schema

```
CREATE OR REPLACE VIEW referencemodel."MSRL_Zuordnung_Dienste" AS
 SELECT dienste.id AS dienst_id, msrl_deskriptoren.id AS msrl_deskriptor_id, dienste.name
      AS dienst_bez, msrl_deskriptoren.nummer AS msrl_deskriptor_nr, msrl_deskriptoren.
     bedeutung AS msrl_deskriptor_bed
 FROM referencemodel.msrl_zuweisung_dienste, referencemodel.msrl_deskriptoren,
     referencemodel.dienste
 WHERE msrl_zuweisung_dienste.dienst_id = dienste.id AND msrl_zuweisung_dienste.
     msrl_deskriptor_id = msrl_deskriptoren.id;
```

Listing A.1: *SQL statement to define a view showing services, their institutions and MSFD descriptors they belong to*

A.1.2 Registration of data sets and services

```
<html>
 <head>
  <title>Diensteeingabeformular MDI-DE</title>
 <?php
 // Connect to database
 ...
 //Get MSFD descriptors
 $query2 = 'SELECT * FROM referencemodel.msrl_deskriptoren';
 ...

 //Get assignments from services to MSFD descriptors
 $query3 = 'SELECT * FROM referencemodel."MSRL_Zuordnung" WHERE datensatz_id='.$id;
 ...

 //Get INSPIRE themes
 $query4 = 'SELECT * FROM referencemodel.inspire_themen';
 ...

 //Get assignments from services to INSPIRE themes
 $query5 = 'SELECT * FROM referencemodel."INSPIRE_Zuordnung_Dienste" WHERE dienst_id='.
     $id;
 ...
?>
<script language="JavaScript">
  function transferOptions(selectObj,insertObj) {
 insertObjLengthStart=insertObj.length;

 if (insertObjLengthStart==0) bereitsDa=false;
 for(j=0; j<selectObj.length; j++) {
  if (selectObj.options[j].selected) {
   for(i=0; i<insertObj.length; i++) {
    if (insertObj.options[i].value!=selectObj.options[j].value){
     bereitsDa=false;
```

```
33      }
34      else {
35       bereitsDa=true;
36      }
37     }
38    }
39   }
40   if (bereitsDa==false){
41    addOptions(selectObj,insertObj);
42    substractOptions(selectObj);
43   }
44   else {
45    alert('Nummer befindet sich bereits dort!');
46   }
47   if (insertObjLengthStart==insertObj.length) {
48     alert('Erst Schlagwörter im rechten Feld auswählen!');
49   }
50   }
51
52    function addOptions(selectObj,insertObj) {
53      ...
54
55    function substractOptions(selectObj) {
56      ...
57  </script>
58  </head>
59  <body onload="document.formular.bez.focus();">
60  <h2>Diensteeingabeformular MDI-DE</h2>
61  <form action="action.php" method="post" name="formular">
62  <fieldset>
63   <legend>Allgemeine Angaben zum Dienst</legend>
64   <p><label for="bez">Bezeichnung:</label> <input type="text" name="bez" /></p>
65   <p><label for="sysname">Systemname:</label> <input type="text" name="sysname" /></p>
66   <p><label for="ores">Onlineresource:</label> <input size="100" type="text" name="ores" />
       </p>
67   <p><label for="grp">Gruppe:</label> <input type="text" name="grp" /></p>
68    <p><label for="thema">Themenzugehörigkeit:</label>
69    <select name="thema">
70     <?
71     for($i=0; $i < $themen_length; $i++) {
72      if ($thema[0]==$thema_id[$i]) echo "<option value=\"$thema_id[$i]\" selected>
          $thema_thema[$i]</option>";
73      else echo "<option value=\"$thema_id[$i]\"> $thema_thema[$i]</option>";
74      }
75     ?>
76          </select>
77   </p>
78    <p><label for="grp">Metadaten vorhanden?</label><input type="radio" name="metadaten"
        value="true">ja<input type="radio" name="metadaten" value="false">nein</p>
79     <p><label for="grp">Metadaten Typ:</label> <input type="text" name="metadaten_typ" /><
         /p>
80   <p><label for="org">Organisation:</label>
81   <br>
82   <?
83    for($i=0; $i < $s; $i++) {
84     echo "<span class=\"radiotrenner\"><input type=\"radio\" name=\"org\" value=\"$orgIds[
         $i]\"> $org[$i]</span>";
85     }
86   ?>
87   </p>
88  </fieldset>
89  <fieldset>
90   <legend>OGC-spezifische Angaben zum Dienst</legend>
91   <p><label for="typ">Servicetyp:</label> <input type="text" name="typ" /></p>
```

```
 <p><label for="vers">Serviceversion:</label> <input type="text" name="vers" /></p>
</fieldset>
<fieldset>
 <legend>Berichte-spezifische Angaben zum Dienst</legend>
...
 <label for="insp">INSPIRE-Themen:</label>
  <table border="0" cellspacing="0" cellpadding="0">
 <tr valign="top">
  <td>
    <select id="links_insp" name="insp_selected[]" size="8" multiple="multiple" style="
        width:160px">
    <?
    for($i=0; $i < $insp_datensatz_length; $i++){
     echo '<option title='.str_replace(' ', ' ', htmlentities(
         $inspire_themen_bedeutung_datensatz[$i])).' value='.$inspire_themen_id_datensatz
         [$i].'>Annex'.$inspire_annex_nr_datensatz[$i].'-Thema'.
         $inspire_themen_nr_datensatz[$i].'</option>';
      }
    ?>
    </select>
  </td>
  <td align="center" valign="middle" width="1">
    <input type="button" name="transferRL" value="&lt;&lt;" onClick=transferOptions(
        document.getElementById('rechts_insp'),document.getElementById('links_insp'))>
    <input type="button" name="transferLR" value="&gt;&gt;" onClick=transferOptions(
        document.getElementById('links_insp'),document.getElementById('rechts_insp'))>
  </td>
  <td>
    <select id="rechts_insp" name="inps_all" size="8" multiple="multiple" style="width
        :160px">
    <?
    for($i=0; $i < $insp_length; $i++){
     echo '<option title='.str_replace(' ', ' ', htmlentities($insp_bedeutung[$i]))
         .' value='.$insp_id[$i].'>Annex'.$insp_annex[$i].'-Thema'.$insp_nummer[$i].'</
         option>';
      }
    ?>
    </select>
  </td>
 </tr>
  </table>
...
```

Listing A.2: *Form to register a service into MDI-DE database*

```
<?php
 if ($id=='') {
  $query2 = "INSERT INTO referencemodel.dienste (servicetyp, serviceversion, name,
      bereitstellende_organisation_id, id, beschreibung, entwicklungsstatus, software,
      onlineresource, systemname, bemerkungen, kategorie, themenzugehoerigkeit, temp,
      quelle_id, gruppe, author, date, metadaten, metadaten_typ) VALUES ('$_POST[typ]', '
      $_POST[vers]', '$_POST[bez]', '$_POST[org]', '$id_new', '$_POST[beschr]', '', '', '
      $_POST[ores]', '$_POST[sysname]', '$_POST[bemerk]', '', '$_POST[thema]', '0', '0',
      '$_POST[grp]', '$_POST[author]', CURRENT_DATE, '$_POST[metadaten]', '$_POST[
      metadaten_typ]')";
 }
 else $query2 = "UPDATE referencemodel.dienste SET servicetyp = '$_POST[typ]',
     serviceversion = '$_POST[vers]', name = '$_POST[bez]',
     bereitstellende_organisation_id = '$_POST[org]', id = '$id', beschreibung = '$_POST[
     beschr]', entwicklungsstatus = '', software = '', onlineresource = '$_POST[ores]',
     systemname = '$_POST[sysname]', bemerkungen = '$_POST[bemerk]', kategorie = '',
     themenzugehoerigkeit = '$_POST[thema]', temp = '0', quelle_id = '0', gruppe = '
     $_POST[grp]', author = '$_POST[author]', date = CURRENT_DATE, metadaten = '$_POST[
     metadaten]', metadaten_typ = '$_POST[metadaten_typ]' WHERE id='$id'";
```

```
6
7  pg_query($query2) or die ('Einfügen fehlgeschlagen: ' . pg_last_error());
8  $test1=$_POST['msrl_selected'];
9
10 //Add/delete MSFD relations
11 if ($id=='') {
12  if ($test1){
13   foreach ($test1 as $t1) {
14   $query2_msrl = "INSERT INTO referencemodel.msrl_zuweisung_dienste (id, dienst_id,
         msrl_deskriptor_id) VALUES ('$id_msrl', '$id_new', '$t1')";
15   pg_query($query2_msrl) or die ('Einfügen fehlgeschlagen: ' . pg_last_error());
16   $id_msrl = $id_msrl + 1;
17   }
18  }
19 }
20 else {
21  if (count($test1)!='0'){
22   //Delete relations
23   $query2_msrl = "DELETE FROM referencemodel.msrl_zuweisung_dienste WHERE dienst_id = '
         $id'";
24   pg_query($query2_msrl) or die ('Einfügen fehlgeschlagen: ' . pg_last_error());
25   //Add new ones
26   foreach ($test1 as $t1) {
27    echo 'Integer gefunden!!!';
28    $query2_msrl = "INSERT INTO referencemodel.msrl_zuweisung_dienste (id, dienst_id,
          msrl_deskriptor_id) VALUES ('$id_msrl', '$id', '$t1')";
29    pg_query($query2_msrl) or die ('Einfügen fehlgeschlagen: ' . pg_last_error());
30    $id_msrl = $id_msrl + 1;
31   }
32   }
33  else {
34   $query2_msrl = "DELETE FROM referencemodel.msrl_zuweisung_dienste WHERE dienst_id = '
         $id'";
35   pg_query($query2_msrl) or die ('Einfügen fehlgeschlagen: ' . pg_last_error());
36  }
37 }
38 //Add/delete INSPIRE relations
39 ...
```

Listing A.3: *Script to insert newly registered/changed services into MDI-DE database*

A.1.3 Presentation of data sets and services

```
<html>
 <head>
  <title>Web Map Services in MDI-DE</title>
  <script type="text/javascript"><!--
  <?php
  // Connect to database
...
   //toggle id to hide/show entries
   function toggle(id){
     var img = 'img_' + id;
     if(document.getElementById(id).style.display == 'block') {
      document.getElementById(id).style.display = 'none';
      document.getElementById(img).src = 'img/plus.gif';
     } else {
      document.getElementById(id).style.display = 'block';
      document.getElementById(img).src = 'img/minus.gif';
     }
    }
  --></script>
 </head>
 <body>
  <form action="WMSBrowser/WMSBrowser.php" method="post">
   <div id="bla">
     <?
      // Get services
   $result2 = pg_query('SELECT DISTINCT organisation_id FROM referencemodel."
       Akteure_Dienste"') or die('Abfrage fehlgeschlagen: ' . pg_last_error());
   while( $row2=pg_fetch_assoc($result2) ) {
    $orgId[] = $row2['organisation_id'];
   }
   $t = sizeof($orgId);
   $c = 1;
   for($k=0; $k < $t; $k++) {
    $resultOrg = pg_query('SELECT organisation_name, gruppe, name, id, onlineresource
        FROM referencemodel."Akteure_Dienste" WHERE organisation_id='.$orgId[$k]) or die
        ('Abfrage fehlgeschlagen: ' . pg_last_error());
    unset($row3);
    unset($grp2);
    unset($name2);
    unset($id2);
    unset($org2);
    unset($ores2);
    while( $row3=pg_fetch_assoc($resultOrg) )
      {
        $grp2[] = $row3['gruppe'];
        $name2[] = $row3['name'];
        $id2[] = $row3['id'];
        $org2[] = $row3['organisation_name'];
        $ores2[] = $row3['onlineresource'];
      }
     $q = sizeof($grp2);

     // build table
     ?><b><?echo $org2[0];?></b><?
     echo "<br>";
     for($m=0; $m < $q; $m++) {
      if ($grp2[$m]!=''){
       ?><a href="#" onclick="toggle('layer<? echo $c ?>')"><img src="img/plus.gif" id="
           img_layer<? echo $c ?>"  border="0"></a> <a href="<?php echo '/MDI-DE/
           dienst_test.php?id='.$id2[$m]; ?>" target="_blank"><?php echo $grp2[$m].' - '.
           $name2[$m]; ?></a> - <i><a href="<?php echo '/MDI-DE/WMSBrowser/WMSBrowser.php
           ?id='.$id2[$m]; ?>" target="_blank">Ansicht im WMSBrowser</a>, <a href="<?php
```

```
            echo $ores2[$m].'request=GetCapabilities&service=WMS'; ?>" target="_blank">
            Capabilities</a>, <a href="<?php echo '/MDI-DE/wms_parser.php?id='.$id2[$m];
            ?>" target="_blank">Capabilities in menschenlesbarer Form</a><br></i><?
56          unset($result_lay);
57          unset($name_lay);
58          unset($titel);
59          unset($id_lay);
60          $result_lay = pg_query("SELECT titel, name, id FROM referencemodel.layers WHERE
                zugehoeriger_dienst_id='$id2[$m]'") or die('Abfrage fehlgeschlagen: ' .
                pg_last_error());
61          while( $row_lay=pg_fetch_assoc($result_lay) )
62          {
63           $name_lay[] = $row_lay['name'];
64           $titel[] = $row_lay['titel'];
65           //$id_lay[] = $row_lay['id'];
66          }
67          $l_n = sizeof($titel);
68          $divText = '';
69          for($j=0; $j < $l_n; $j++)
70           {
71            if($j!=0) $divText = $divText.'<br><input type="checkbox" name="'.$name_lay[$j
                ].'" value="'.$id2[$m].'">'.$titel[$j];
72            else $divText = '<input type="checkbox" name="'.$name_lay[$j].'" value="'.$id2[
                $m].'">'.$titel[$j];
73           }?>
74          <div id="layer<? echo $c ?>"><? echo $divText ?></div><?php
75          echo "<br>";
76          $c++;
77         }
78         else {
79          ?><a href="#" onclick="toggle('layer<? echo $c ?>')"><img src="img/plus.gif" id="
                img_layer<? echo $c ?>" border="0"></a> <a href="<?php echo '/MDI-DE/
                dienst_test.php?id='.$id2[$m]; ?>" target="_blank"><?php echo $name2[$m]; ?></
                a> - <i><a href="<?php echo '/MDI-DE/WMSBrowser/WMSBrowser.php?id='.$id2[$m];
                ?>" target="_blank">Ansicht im WMSBrowser</a>, <a href="<?php echo $ores2[$m
                ].'request=GetCapabilities&service=WMS'; ?>" target="_blank">Capabilities</a>,
                 <a href="<?php echo '/MDI-DE/wms_parser.php?id='.$id2[$m]; ?>" target="_blank
                ">Capabilities in menschenlesbarer Form</a><br></i><?
80          unset($result_lay);
81          unset($name_lay);
82          unset($titel);
83          unset($id_lay);
84          $result_lay = pg_query("SELECT titel, name, id FROM referencemodel.layers WHERE
                zugehoeriger_dienst_id='$id2[$m]'") or die('Abfrage fehlgeschlagen: ' .
                pg_last_error());
85          while( $row_lay=pg_fetch_assoc($result_lay) )
86          {
87           $name_lay[] = $row_lay['name'];
88           $titel[] = $row_lay['titel'];
89           //$id_lay[] = $row_lay['id'];
90          }
91          $l_n = sizeof($titel);
92          $divText = '';
93          for($j=0; $j < $l_n; $j++)
94           {
95            if($j!=0) $divText = $divText.'<br><input type="checkbox" name="'.$name_lay[$j
                ].'" value="'.$id2[$m].'">'.$titel[$j];
96            else $divText = '<input type="checkbox" name="'.$name_lay[$j].'" value="'.$id2[
                $m].'">'.$titel[$j];
97           }?>
98          <div id="layer<? echo $c ?>"><? echo $divText ?></div><?php
99          echo "<br>";
100         $c++;
101        }
```

```
102       }
103       echo "<hr>";
104     }
105     ?>
106   </div>
107 ...
```

Listing A.4: *Overview of all services in MDI-DE (Hide/Show layers, Capabilities etc.)*

```
1  <html>
2   <head>
3    <title>WMSBrowser</title>
4     <script type="text/javascript" src="http://extjs.cachefly.net/ext-3.2.1/adapter/ext/
          ext-base.js"></script>
5     <script type="text/javascript" src="http://extjs.cachefly.net/ext-3.2.1/ext-all.js"></
          script>
6     <script src="http://www.openlayers.org/api/2.10/OpenLayers.js"></script>
7     <script src="lib/GeoExt.js" type="text/javascript"></script>
8     <script src="lib/GeoExt/SingleFile.js"></script>
9     <script src="lib/GeoExt.ux/widgets/WMSBrowser.js"></script>
10    <link rel="stylesheet" type="text/css" href="css/ext-all.css" />
11    <link rel="stylesheet" type="text/css" href="css/examples.css" />
12    <link rel="stylesheet" type="text/css" href="css/WMSBrowser.css" />
13    <script src="http://www.openstreetmap.org/openlayers/OpenStreetMap.js"></script>
14    <script type="text/javascript"><!--
15    var WMSBrowser, mapPanel, browserWindow, tree;
16
17    <?php
18     //connect to database
19     ...
20     ?>var ores = new Array();<?
21     $j = 0;
22     $s = sizeof($onlineres);
23     for($i=0; $i < $s; $i++)
24        {
25      if($onlineres[$i]!=''){
26        ?>ores[<? echo "$j"; ?>] = new Array();
27        ores[<? echo "$j"; ?>].push("<? echo "$onlineres[$i]"; ?>");<?
28        $j++;
29       }
30     }
31     ?>var layer = new Array();<?
32     if ($id!=''){
33      unset($onlineres);
34
35      $query2 = 'SELECT * FROM referencemodel."Layers_Dienste_Orgs" WHERE dienst_id='.$id;
36      $result2 = pg_query($query2) or die('Abfrage fehlgeschlagen: ' . pg_last_error());
37
38      while( $row2=pg_fetch_assoc($result2) )
39      {
40        $onlineres[] = $row2['dienst_onlineresource'];
41        $titel[] = $row2['layer_titel'];
42        $name[] = $row2['layer_name'];
43        $grp_rsrc=array_merge($onlineres, $titel);
44        $ressourcen=array_merge($grp_rsrc, $name);
45      }
46      $s2 = sizeof($onlineres);
47
48      for($i=0; $i < $s2; $i++)
49         {
50        ?>layer[<? echo "$i"; ?>] = new Object();
51        layer[<? echo "$i"; ?>]["Resource"] = "<? echo "$ressourcen[$i]"; ?>";
52        layer[<? echo "$i"; ?>]["Titel"] = "<? echo "$titel[$i]"; ?>";
53        layer[<? echo "$i"; ?>]["Name"] = "<? echo "$name[$i]"; ?>";<?
```

```
    }
   }
   else if (sizeof($layer)>0) {
    for($j=0; $j < sizeof($layer); $j++) {
     $dienst_id = $_POST[$layer[$j]];
     $name_layer = $layer[$j];
     unset($result_lay);
     unset($name_lay);
     unset($titel);
     unset($query2);
     unset($result2);
     unset($row2);
     unset($onlineres);

     $query2 = 'SELECT dienst_onlineresource FROM referencemodel."Layers_Dienste_Orgs"
         WHERE dienst_id='.$dienst_id;
     $result2 = pg_query($query2) or die('Abfrage fehlgeschlagen: ' . pg_last_error());

     while( $row2=pg_fetch_assoc($result2) )
     {
       $onlineres[] = $row2['dienst_onlineresource'];
     }

     $result_lay = pg_query("SELECT titel, name, id FROM referencemodel.layers WHERE
         zugehoeriger_dienst_id='$dienst_id' AND name='$name_layer'") or die('Abfrage
         fehlgeschlagen: ' . pg_last_error());
     while( $row_lay=pg_fetch_assoc($result_lay) ) {
      $name_lay[] = $row_lay['name'];
      $titel[] = $row_lay['titel'];
     }
     ?>layer[<? echo "$j"; ?>] = new Object();
     layer[<? echo "$j"; ?>]["Resource"] = "<? echo "$onlineres[0]"; ?>";
     layer[<? echo "$j"; ?>]["Titel"] = "<? echo "$titel[0]"; ?>";
     layer[<? echo "$j"; ?>]["Name"] = "<? echo "$name_lay[0]"; ?>";<?
    }
   }
  ?>
    var oServerStore  = new Ext.data.SimpleStore({
    fields: ['url'],
    data : ores
   });

   function createLayer(map, layer, resource, int_name) {
    var wms = new OpenLayers.Layer.WMS(int_name, resource, {layers: layer, format: 'image
        /png',transparent: true }, {isBaseLayer: false});
    map.addLayer(wms);
   }

   Ext.onReady(function() {
    Ext.QuickTips.init();

    var options = {
     projection: "EPSG:4326",
     'numZoomLevels': 20
    };

    map = new OpenLayers.Map('map', options);

    mapPanel = new GeoExt.MapPanel({
     region: "center",
     layers: [new OpenLayers.Layer.WMS("OSM",
      "http://labs.metacarta.com/wms/vmap0",
      {layers: "basic"})] ,
     center: [11,55],
```

```
     zoom: 6,
     map: map
    });

    if (layer.length>0){
     for (var i = 0; i < layer.length; i++) {
      createLayer(map, layer[i]["Name"], layer[i]["Resource"], layer[i]["Titel"]);
     }
    }

    WMSBrowser = new GeoExt.ux.WMSBrowser({
     title: OpenLayers.i18n("WMSBrowser"),
     region: "east",
     width: 500,
     gridPanelOptions: {'height': 250},
     allowInvalidUrl: true,
     serverStore: oServerStore,
     layerStore: mapPanel.layers
    });
...
```

Listing A.5: *WMSBrowser to visualize selected services*

A.2 SSCVisualizer

```
function getResults() {
 selServices = $('#services').val();
 if (selServices.length == 'null') $('#console').html("<b>Error:</b> No service selected
     ");
 var scores = [];
 var dates = [];
 var scoreOutput="<ul>";
 for (var i in jsondata.data) {
  for (var j in selServices) {
   //Check if score for a date already exists (there 2 measurements per day)
   //the array scores is defined and has at least one element
   if (typeof scores !== 'undefined' && scores.length > 0) {
    for (var k in scores) {
     var score = [];
     //21600000 = 6h hinzufügen beim 2. Wert
     if(scores[k][0][0] == removeTime(jsondata.data[i].date).replaceAll('-','')) {
      if (jsondata.data[i].summary.scoredTest.currentScore == "-99.99") score.push([
          parseDate(jsondata.data[i].date).getTime()+21600000, 0, jsondata.data[i].id]);
      else score.push([parseDate(jsondata.data[i].date).getTime()+21600000, jsondata.
          data[i].summary.scoredTest.currentScore, jsondata.data[i].id, jsondata.data[i
          ].name, jsondata.data[i].summary.scoredTest.currentSpeed]);
     }
     //date+time is new
     else {
      if (jsondata.data[i].summary.scoredTest.currentScore == "-99.99") score.push([
          parseDate(jsondata.data[i].date).getTime(), 0, jsondata.data[i].id]);
      else score.push([parseDate(jsondata.data[i].date).getTime(), jsondata.data[i].
          summary.scoredTest.currentScore, jsondata.data[i].id, jsondata.data[i].name,
          jsondata.data[i].summary.scoredTest.currentSpeed]);

     }
    }
   }
   //initially set up the scores array and fill it with the first object
   else {
    var score = [];
    if (jsondata.data[i].summary.scoredTest.currentScore == "-99.99") score.push([
        parseDate(jsondata.data[i].date).getTime(), 0, jsondata.data[i].id]);
```

```
31      else score.push([parseDate(jsondata.data[i].date).getTime(), jsondata.data[i].
            summary.scoredTest.currentScore, jsondata.data[i].id, jsondata.data[i].name,
            jsondata.data[i].summary.scoredTest.currentSpeed]);
32     }
33     if (jsondata.data[i].id == selServices[j]) {
34      dates.push(removeTime(jsondata.data[i].date));
35      scoreOutput+="<li>" +  jsondata.data[i].id + " - " + parseDate(jsondata.data[i].date
            ).getTime() + " - " + jsondata.data[i].summary.scoredTest.currentScore + "</li>"
            ;
36      scores.push(score);
37     }
38    }
39    if (typeof scores !== 'undefined' && scores.length > 0) {
40    }
41    scoreOutput+="</ul>";
42   }
43   return scores;
44  }
```

Listing A.6: *Function that returns the dates, scores, id and names for selected services*

```
1   function transformScores(scores, type) {
2    var ids = [];
3    var idtemp = "empty";
4    var ds = new Array();
5
6    var barWidth = calcBarWidth();
7
8    for (var i = 0, j = scores.length; i < j; i++) {
9     if (typeof(ids[0]) == 'undefined') {
10     //fill ids[] with first id
11     ids.push(scores[i][0][2]);
12     //create and fill temp[] with 1st score
13     var temp = [];
14     if (type=='score') temp.push([scores[i][0][0], scores[i][0][1]]);
15     if (type=='speed') temp.push([scores[i][0][0], scores[i][0][4]]);
16    }
17    //2. case: there is something in ids[]
18    else {
19     for (var k in ids) {
20      if(scores[i][0][2] == ids[k]) {
21       //found an existing id -> add score to temp
22       if (type=='score') temp.push([scores[i][0][0], scores[i][0][1]]);
23       if (type=='speed') temp.push([scores[i][0][0], scores[i][0][4]]);
24      }
25      else if (idtemp !== scores[i][0][2]){
26       //found a new id -> put new id in ids[]
27       ids.push(scores[i][0][2]);
28       //push old scoreData in temp into ds[];
29       ds.push({
30        label:scores[i-1][0][3],
31        data:temp,
32        bars: {
33         show: true,
34         barWidth: calcBarWidth(),
35         order: 1,
36         lineWidth : 1
37        }
38       });
39       //create and fill temp[] with 1st score for the new id
40       var temp = [];
41       if (type=='score') temp.push([scores[i][0][0], scores[i][0][1]]);
42       if (type=='speed') temp.push([scores[i][0][0], scores[i][0][4]]);
```

```
      //save last new found id so the in future loops it does not get recognized as a new
          id
      idtemp = scores[i][0][2];
     }
    }
   }
   //3. case: last run of the outer for loop -> temp[] has to be pushed into ds[]
   if (typeof(scores[i+1]) == 'undefined') {
    //push old scoreData in temp into ds[];
    ds.push({
     label:scores[i-1][0][3],
     data:temp,
     bars: {
      show: true,
      barWidth: calcBarWidth(),
      order: 1,
      lineWidth : 1
     }
    });
   }
  }
  return ds;
}
```

Listing A.7: *Function that transforms an array into an array suitable for Flot*

```
function saveDiagram() {
 if (AggrOrComplex=='aggregate') {
 //Save the diagram only
  html2canvas(document.getElementById("scoreGraph"), {
   onrendered: function(canvas) {
     Canvas2Image.saveAsPNG(canvas);
    }
   });
 }
 if (AggrOrComplex=='complex') {
 //Saving the diagram AND the legend in one image
  var legendCanvas;
  var diagramCanvas;
  html2canvas(document.getElementById("legend"), {
   onrendered: function(canvas) {
    legendCanvas = canvas;
    html2canvas(document.getElementById("scoreGraph"), {
     onrendered: function(canvas) {
      diagramCanvas = canvas;

      //Commbine the two images (legend+diagram), source: http://jsfiddle.net/m1erickson
          /5JTtd/
      var diagramAndLegendCanvas=document.createElement("canvas");
      var ctx=diagramAndLegendCanvas.getContext("2d");
      diagramAndLegendCanvas.width=diagramCanvas.width;
      diagramAndLegendCanvas.height=diagramCanvas.height;

      //draw all 2 images into 1 combined image
      ctx.drawImage(diagramCanvas,0,0);
      ctx.drawImage(legendCanvas,diagramCanvas.width-legendCanvas.width,0);

      //Save
      Canvas2Image.saveAsPNG(diagramAndLegendCanvas);
     }
    });
   }
  });
 }
```

```
}
```

Listing A.8: *Function saveDiagram()*

```
<html>
<head>
 <meta http-equiv="Content-Type" content="text/html;charset=utf-8" />
 <title>SSCVisualizer</title>
  <script src="http://code.jquery.com/jquery-1.9.1.js"></script>
 <script language="javascript" type="text/javascript" src="http://www.flotcharts.org/flot/jquery.flot.js"></script>
 <script language="javascript" type="text/javascript" src="http://nihilogic.dk/labs/canvas2image/canvas2image.js"></script>
 <script language="javascript" type="text/javascript" src="http://html2canvas.hertzen.com/build/html2canvas.js"></script>
 <script src="sscv_new.js"></script>
</head>
<body>
 <h1>SSCVisualizer</h1>
 <i>Hint: Use Mousewheel to zoom and click'n'hold mouse button to pan</i><p>
 <div id="main">
  <div id="input">
   <h2>Input</h2>
   <section id="fetch">
     Authentification: <input type="text" placeholder="Enter auth id" id="auth" style="width:240px;"></input>
     Type:
    <select id="type">
     <option value="WMS">WMS</option><option value="WFS">WFS</option>
    </select>
     <br>
      <input type="radio" id="period1" name="period" value="7"> last week<br>
      <input type="radio" id="period2" name="period" value="14"> last two weeks<br>
      <input type="radio" id="period3" name="period" value="9999"> all (beware!)<br>
      <input type="radio" id="period4" name="period"> set date range
     </section>
   <label for="from">From</label>
   <input type="text" id="from" name="from" onclick="checkRadio()"/>
   <br>
   <label for="to">to</label>
   <input type="text" id="to" name="to"  onclick="checkRadio()"/>
   <div id="reqLink"></div>
   <input class="buttons" type="button" value="Fetch JSON & get Services" id="FetchBtn" onClick="jsonptest()"/>
   <div id="serviceList"></div>
   <div id="buttonArea">
    <h2>Action</h2>
     <input class="buttons" type="button" value="Show Score" id="ScoreBtn" onClick="showResults('score')" disabled="true"/>
     <input class="buttons" type="button" value="Show Speed" id="SpeedBtn" onClick="showResults('speed')" disabled="true"/>
     <input class="buttons" type="button" value="Aggr. Scores" id="ScoreAgBtn" onClick="aggregate('score')"/>
     <input class="buttons" type="button" value="Aggr. Speeds" id="SpeedAgBtn" onClick="aggregate('speed')"/>
     <input class="buttons" type="button" value="Save Diagram as PNG" id="SaveBtn" onClick="saveDiagram()" disabled="true"/>
     <input class="buttons" type="button" value="Test" id="TestBtn" onClick="test()"/>
   </div>
    </div>

  <div id="legendArea">
   <h2>Legend</h2>
   <div id="legend"></div>
```

```
  </div>
    <div id="scoreGraph">
  <h2>Diagram</h2>
    </div>
 </div>

 <div class="console">
  <b>Console</b>
   <div id="console">
  </div>
 </div>
 <hr>
</body>
</html>
```

Listing A.9: *SSCVisualizer HTML code*

A.3 JSKOSify

A.3.1 Main class

```
/**
 * Builds a new element for the upcoming concept with the outcommings of the tableArray
     and the cell content
 * @param elemTerm term in the column "header" in the current row of the CSV file (a
     single cell in fact)
 * @param skosCorrespondent the SKOS element to use (e.g. prefLabel, altLabel,
     definition etc.)
 * @param language the language of the term
 * @return an {@link Element} ready to be added by the concept
 */
private static Element buildElem(String elemTerm, String skosCorrespondent, String
    language) {
 //Check if there is content at all
 if (!elemTerm.equals("") && !skosCorrespondent.equals("")){
  //Remove the "skos:" in front of the SKOS element (eg "skos:prefLabel" becomes just "
      prefLabel"
  String skosElement = skosCorrespondent.substring(skosCorrespondent.indexOf(":")+1,
      skosCorrespondent.length());
  Element newElem = new Element(skosElement,nsSKOS);
  newElem.setAttribute("lang",language,org.jdom2.Namespace.XML_NAMESPACE);
  newElem.setText(elemTerm);

  return newElem;
  }
 else return null;
}
```

Listing A.10: *Function buildElem()*

```
/**
 * If a concept has (a) broader concept(s) (e.g. beach scarp has beach as broader
     concept) this is added to the concept as skos:broader.
 * Also, the superordinate concept is deleted so that the narrower concept (e.g. beach
     scrap) is display directly under the broader concept (here: beach) in the tree view
 * @param doc the Document to be altered
 * @param namespace the namespace used in the URI of the concepts
 * @param topTerm the superordinate concept's name which is also used in the URI of each
     concept
 * @param startOnly If true a suffix regex (SQL syntax: term LIKE 'accident%') is used
     for the search, and if false a prefix/suffix combined regex (SQL syntax: term LIKE
     '%accident%') is used for the search
```

```
8   * @return a JDOM {@link Document} with (hopefully) many skos:broader elements added to
        some concepts
9   */
10  public static Document addRelations(Document doc, String namespace, String topTerm,
        boolean startOnly) {
11   int i = 1;
12   //Get all children
13   List<Element> children = doc.getRootElement().getChildren();
14   int childrenNum = children.size();
15   //check if there are children
16   if (children.size() > 0) {
17    //iterate through the children
18    Iterator<Element> iter = children.iterator();
19    while(iter.hasNext()) {
20     //Get the current concept/element
21        Element current = (Element)iter.next();
22
23        //Get the URI of the current concept
24        String conceptURI = "";
25        if (current.getAttribute("about", nsRDF)!=null) conceptURI = current.
              getAttributeValue("about", nsRDF);
26
27     //Check if there is a broader term for the term handled at the moment
28     List<Element> broaderElems = checkForBroaderTerms(doc, conceptURI, current, namespace
           , topTerm, startOnly);
29
30     if (!broaderElems.isEmpty()) {
31      //If there is/are broader term/s first delete the existing one (so that it shows
            directly under the broader term in the tree)
32      current.removeChild("broader", nsSKOS);
33      //Add the broader term/s
34      for (Element broaderElem : broaderElems) current.addContent(broaderElem);
35     }
36     i++;
37    }
38   }
39   return doc;
40  }
```

Listing A.11: *Function addRelations()*

```
1   /**
2    * Saves a JDOM {@link Document} as {@link RDF} in the given {@link RDFFormat}
3    * @param doc the JDOM {@link Document} to be saved as {@link RDF}
4    * @param namespace the namespace to resolve any relative URI references (if everything
         worked out as it should this could be "" or anything, too)
5    * @param rdfFilePath the path where the RDF file will be saved
6    * @param formatRDF The {@link RDFFormat} the file will be saved in
7    */
8   public static void saveRDF(Document doc, String namespace, String rdfFilePath, String
        formatRDF) throws FileNotFoundException {
9    //initialize a RDFWriter
10   RDFWriter rdfWriter = null;
11   //initialize a XMLOutputter
12   XMLOutputter xmlOutput = new XMLOutputter();
13   Format format = Format.getPrettyFormat();
14   //because of German language set encoding to ISO-8859-1
15   format.setEncoding(encoding);
16   xmlOutput.setFormat(format);
17   //Output the JDOM Document as a String
18   String xmlStr = xmlOutput.outputString(doc);
19
20   //initialize a StringReader with the given String (Document) because that is what the
         RDFParser is able to parse
```

```
21   StringReader strReader = new StringReader(xmlStr);
22
23   //Set the right RDFWriter for the format specified
24   if (formatRDF.equals("NTriples")) rdfWriter = new NTriplesWriter(new FileOutputStream (
         rdfFilePath));
25   else if  (formatRDF.equals("Turtle")) rdfWriter = new TurtleWriter(new FileOutputStream
          (rdfFilePath));
26   else if  (formatRDF.equals("RDFXML")) rdfWriter = new RDFXMLWriter(new FileOutputStream
          (rdfFilePath));
27
28   //initialize a RDFParser
29   RDFParser rdfParser = new RDFXMLParser();
30   //and set the RDFWriter as its RDFHandler
31   rdfParser.setRDFHandler(rdfWriter);
32   try {
33    //parse the String (Document) into the right format
34    rdfParser.parse(strReader, namespace);
35   } catch (RDFParseException | RDFHandlerException | IOException e) {
36    // TODO Auto-generated catch block
37    e.printStackTrace();
38   }
39    }
```

Listing A.12: *Function saveRDF()*

A.3.2 GUI class

```
1  private void createTable() {
2   table = new JTable();
3
4   Object[][] data = {};
5
6   String[] columnNames = {column1, column2, column3, column4, column5, column6};
7
8   table.setModel(new DefaultTableModel(data, columnNames));
9
10  //Add hierarchy level ComboBoxes to the fourth column
11  addComboboxToTable(3, new String[]{"","1", "2", "3", "4", "5", "6", "7", "8", "9", "10"
        });
12
13  //Add Language ComboBoxes to the third column
14  addComboboxToTable(2, new String[]{"","German", "English", "French"});
15
16  //Add skos correspondents to the second column
17  addComboboxToTable(1, new String[]{"skos:prefLabel", "skos:altLabel", "skos:
        editorialNote", "skos:definition", "skos:broader", "skos:narrower"});
18
19  table.setRowHeight(25);
20
21  table.getColumnModel().getColumn(4).setCellRenderer(table.getDefaultRenderer(Boolean.
        class));
22  table.getColumnModel().getColumn(4).setCellEditor(table.getDefaultEditor(Boolean.class)
        );
23
24  table.getColumnModel().getColumn(5).setCellRenderer(table.getDefaultRenderer(Boolean.
        class));
25  table.getColumnModel().getColumn(5).setCellEditor(table.getDefaultEditor(Boolean.class)
        );
26
27  JScrollPane scrollPane = new JScrollPane(table);
28  frmTest.getContentPane().add(scrollPane, BorderLayout.CENTER);
29 }
```

```
30
31  private void addComboboxToTable(int vColIndex, String[] values) {
32   DefaultTableModel model = (DefaultTableModel)table.getModel();
33
34   TableColumn col = table.getColumnModel().getColumn(vColIndex);
35   col.setCellEditor(new MyComboBoxEditor(values));
36
37   // If the cell should appear like a combobox in its
38   // non-editing state, also set the combobox renderer
39   col.setCellRenderer(new MyComboBoxRenderer(values));
40  }
```

Listing A.13: *Functions createTable() and addComboboxToTable()*

A.3.3 Tools package

```
1  public class GemetFinder {
2   private static String getURL(String getMethod, String searchTerm, String lang, String
        domain, String thesaurus_uri, int search_mode) {
3    String searchURL = "";
4    if (searchTerm!=null && lang!=null && thesaurus_uri!=null) {
5     if (!searchTerm.equals("") && !lang.equals("") && !thesaurus_uri.equals("") &&
          search_mode>=0 && search_mode<=4) {
6      searchURL = domain
7         +getMethod
8         +"?thesaurus_uri="+thesaurus_uri
9         +"&language="+lang
10        +"&search_mode="+search_mode
11        +"&keyword="+searchTerm.replace(" ", "+").replace("\n", "+").replace("\r", "+");
12     return searchURL;
13    }
14    else return null;
15   }
16   else return null;
17  }
18
19  private static String readAll(BufferedReader buffReader) throws IOException {
20         StringBuilder stringBuilder = new StringBuilder();
21         int i;
22         while ((i = buffReader.read()) != -1) {
23          stringBuilder.append((char) i);
24         }
25         return stringBuilder.toString();
26       }
27
28     private static JSONArray readJsonFromUrl(String url) throws IOException,
           JSONException {
29         InputStream inputStream = new URL(url).openStream();
30         try {
31           BufferedReader buffReader = new BufferedReader(new InputStreamReader(
                 inputStream, Charset.forName("UTF-8")));
32           String jsonText = readAll(buffReader);
33           JSONArray json = new JSONArray(jsonText);
34           return json;
35         } finally {
36         inputStream.close();
37         }
38       }
39
40     public static Element getCloseMatchElem(String getMethod, String searchTerm, String
           lang, String domain, String thesaurus_uri, int search_mode) throws IOException,
           JSONException {
```

```
41      Element closeMatch;
42      String gemetResult = findTermReturn1stURI(getMethod, searchTerm, lang, domain,
            thesaurus_uri, search_mode);
43
44    if (gemetResult!=null){
45     closeMatch = new Element("closeMatch",CSV2SKOS.nsSKOS);
46     closeMatch.setAttribute("resource",gemetResult,CSV2SKOS.nsRDF);
47     return closeMatch;
48     }
49       else return null;
50      }
51
52   public static String findTermReturn1stURI(String getMethod, String searchTerm, String
        lang, String domain, String thesaurus_uri, int search_mode) throws IOException,
        JSONException {
53    String searchURL = getURL(getMethod, searchTerm, lang, domain, thesaurus_uri,
          search_mode);
54    String resultURI = "";
55    JSONArray json;
56
57    if (searchURL!=null) {
58     json = readJsonFromUrl(searchURL);
59     if (json.length()>0) resultURI = (String) json.getJSONObject(0).get("uri");
60     else resultURI = null;
61    }
62    else resultURI = null;
63
64       return resultURI;
65   }
66  }
```

Listing A.14: *Class GemetFinder*

B In-depth evaluations of MSDIs

B.1 Australia: ASDI, AMSIS and Ocean Portal

A – Data

A1 – Core datasets The IMOS Ocean Portal offers data in the fields multi-decadal ocean change, climate variability and weather extremes, major boundary currents and inter-basin flows, continental shelf processes and biological responses (Moltmann et al., 2010, p. 2) while AMSIS offers the following core datasets (Nairn, 2009, p. 22):

- Administrative Boundaries
  - Maritime limits (Coastal waters, territorial sea, contiguous zone, exclusive economic zone, extended continental shelf, coral sea limits)
  - Territorial sea baselines and basepoints, treaty boundaries
  - Petroleum and submerged land act boundary
  - Offshore petroleum lease and offshore mineral lease boundaries
  - Great Barrier Reef Marine Park boundary and planning zone boundaries
  - World heritage areas
  - Commonwealth marine protected areas, National Estate areas, fisheries
  - Indigeneous land use boundaries (agreements, native title determinations, native title applications)
  - Australian search and rescue area boundary
  - Defence firing practice and exercise areas
  - Customs port limits, security port limits, immigration zone
- Framework Data
  - Shoreline , state borders, islands reefs, rocks, cays, shoals, seas
- Bathymetry
  - Bathymetric image, isobaths
- Coastal and Offshore Gazetteer
  - Cultural locations, land features, marine features
- Anthropogenic Features
  - Historic shipwrecks, ocean disposal sites
- Transport
  - Ship reporting locations, derived shipping lanes, ferry routes
- Infrastructure
  - Petroleum wells, platforms, pipelines, submarine cables, navigational aids
- Geology
  - Seafloor features, sedimentary basins, tectonic elements
- Environmental Management
  - Bioregions
  - Marine planning regions

A2 – Degree of involvement of different agencies/institutions AMSIS contains data from agencies and industry sources like the Australian Fisheries Management Authority (AFMA), the Australian Maritime Safety Authority (AMSA), the Australian Hydrographic Service (AHS), the Department of Environment and Conservation (previously CALM) and many others.[1] As outlined in indicator A1 (on page 163) IMOS is a set of nodes – covering Queensland (QIMOS), New South Wales (NSWIMOS), Southern Australia (SAIMOS), Western Australia (WAIMOS), Tasmania and Bass Strait (TASIMOS), and the Bluewater and Climate (BWC) node (see figure B.2) – and facilities such as Argo Australia,

[1] see http://www.ga.gov.au/amsis/datacontributors.jsp

Ships of Opportunity, Australian National Facility for Ocean Gliders, Australian National Mooring Network, Australian Coastal Ocean Radar Network and many more. (Moltmann et al., 2010), (Proctor et al., 2010) With all these facilities and nodes it can be stated that the datasets are coming from a wide range of agencies resp. institutions.

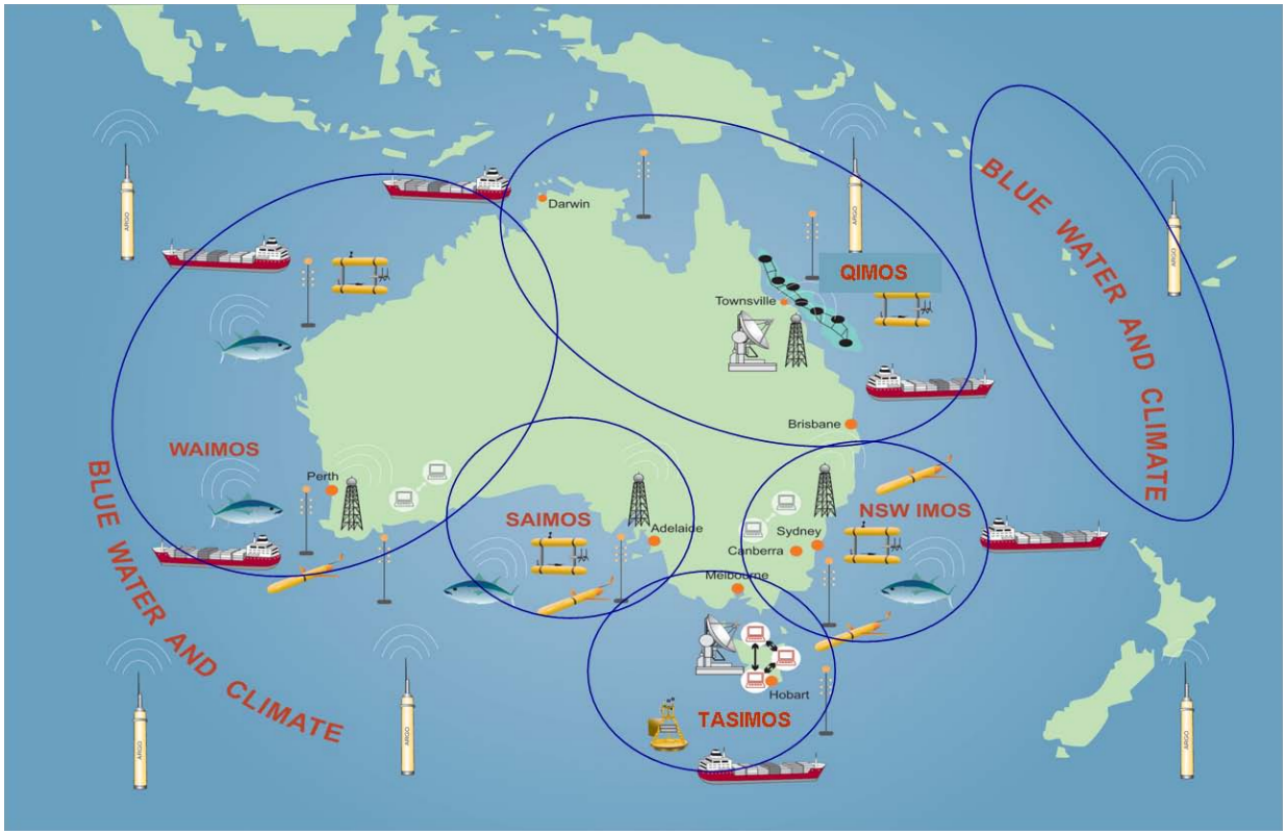


Figure B.2: *IMOS Nodes and Facilities (source: (Proctor et al., 2010, p. 13)*

B – Metadata

B1 - Availability of Metadata/Metadata catalogue (CSW) There is neither a central metadata catalog nor a CSW for AMSIS but you can find all the layers provided in AMSIS through the Australian Spatial Data Directory (ASDD), however, this provides many more datasets through a number of directory systems where each data directory is called a node. The ASDD uses the Z39.50 protocol to simultaneously search the data directories which are based on XML metadata document collections. The underlying Z39.50 servers can also be accessed by Z39.50 client software.[2] IMOS/eMII uses the Metadata Entry and Search Tool (MEST) which is a modified version of the GeoNetwork opensource V2.2.0 metadata catalogue holding ISO 19115/19139 standard records and allowing manual and automatic upload of metadata records (and data sets) and providing data discovery, access and download. A MEST is installed at each node of IMOS which requires a master GeoNetwork catalog (http://imosmest.aodn.org.au) routinely harvesting new metadata records from the regional catalogues so that all IMOS data is kept up to date. (Proctor et al., 2010), (Moltmann et al., 2010) and (de La Beaujardière et al., 2009)

B2 – Data quality and accuracy When we assume that (Nairn, 2009)'s statement from 2009 (" [. . .] increasing use will be made of spatial data infrastructure principles as promoted by ANZLIC such

[2] see http://asdd.ga.gov.au/asdd/tech/architecture.html

as [. . .] metadata management [. . .] ") fulfilled itself and AMSIS relies onto the metadata guidelines of ANZLIC (Australia New Zealand Land Information Council) it can be seen that five metadata elements regarding data quality are included in the core metadata fields: lineage, positional accuracy, attribute accuracy, logical consistency and completeness.
According to (de La Beaujardière et al., 2009) "the IMOS data management infrastructure employs OGC standards wherever possible. The main components of the system are: [. . .] SensorML, which provides standard models and an XML encoding for describing sensors and measurement processes [. . .] ". This means that data quality is built in through sensorML inter alia with its `QualityPropertyType`.

B3 – Coordination (ANZLIC Metadata Working Group, 2001) states that "a mapping between the ANZLIC core metadata elements and the draft version of the ISO 19115 has been prepared" which means that AMSIS which relies on the metadata guidelines of ANZLIC (Australia New Zealand Land Information Council) is using a metadata profile based on the widely adopted metadata standard ISO 19115. According to the install documentation of MEST[3] MEST and thus IMOS "supports the following profiles of the ISO/AS/NZS-19115 geographic metadata standard:

- the ANZLIC profile (version 1.1)
- the Australian Marine Community Profile (MCP) (version 1.3-19139)
- the Australian Defence Organisation (ADO) Profile (version 1.0.3)
- the World Meteorological Organisation (WMO) Core Profile (version 1.0) [. . .]
- the sensorML schema - a gml [sic] application schema for sensors"

In summary it can be said that ANZLIC is coordinating the implementation of metadata guidelines and IMOS in particular relies on well-known international metadata standards.

C – Services and Interfaces

C1 – Availability of Services Because "place names in AMSIS are derived from the Gazetteer of Australia [. . .] "[4] there at least has to be one service which affects AMSIS. According to the Oceans Portal project Governance Working Group discussion paper (Oceans Portal Project Governance Working Group, 2006) " [. . .] the services provided could include: [. . .]

- OGC Web Map Services,
- OGC Web Features Services,
- OGC Web Coverage Services,
- Species Finder Services,
- Name Services,
- Gazetteer Services,
- Modelling Services,
- Download Services,
- Data Access Query Model Services,
- Transformation Mapping Services."

For instance when looking for WMS for the themes *CSIRO Marine and Atmospheric Research Data* and *Australian Institute of Marine Science Data (AIMS)* web map services could be found[5]. Probably many more (even of other service types) could be found.

C2 – Performance Nothing could be found regarding performance of services for AMSIs but in the discussion paper (mentioned in indicator C1) guidelines for availability, reliability and performance of the IMOS Ocean Portal are discussed.

[3] available at: ftp://emii2.its.utas.edu.au/pub/software/IMOS-eMII-MEST/readme-install.html
[4] see http://www.ga.gov.au/amsis/oceangov.jsp#placenames
[5] www.cmar.csiro.au/geoserver/ows?service=wms&version=1.0.0&request=GetCapabilities and http://e-atlas.org.au/maps/ows?SERVICE=WMS&REQUEST=GetCapabilities

C3 – Clearinghouse and geoportal Portals are available for AMSIS (http://www.ga.gov.au/amsis/index.jsp) as well as IMOS Ocean Portal (http://imos.aodn.org.au/webportal/). Both include everything an average geoportal offers from search functionality to map viewer and so on.

C4 – access privileges/Custodianship For AMSIS no precise information could be acquired about the actual role model but (Nairn, 2009) at least states that "the selection of stakeholders was made by researching those that could contribute information [. . .] and agencies that could be potential users." This means that at least two roles were identified for AMSIS. The same is true for IMOS' Ocean Portal. (Moltmann et al., 2010) just lists potential usages "such as management of marine natural resources and their associated ecosystems, support and management of coastal and offshore industries, safety at sea, marine tourism and defence." but states nothing more regarding a role model. But that there are roles inside the Ocean Portal can be seen in figure B.4 (see indicator E2 on page 166). There two roles for the actor *End User* (trusted and anonymous) and furthermore the actors *Administrator* and *Contriubutor* can be found.

D – Standards / D1 – Interoperability Regarding metadata standards for both initiatives please refer to indicator B3 on page 165 and for standards for services see indicator C1 on page 165. Apart from the already mentioned standards IMOS' Ocean Portal uses netCDF (Network Common Data Form) because " [. . .] a significant proportion of IMOS data is of either gridded (satellite, HF radar) or in timeseries form (Argo, ship of opportunity, gliders, moorings, networked sensors) and could sensibly written into a self describing format (netCDF) [. . .] ". (Proctor et al., 2010)

E – Modelling

E1 – existence of a government policy for SDI According to the information found in the AMSIS portal (http://www.ga.gov.au/amsis/oceangov.jsp) AMSIS is strongly connected to Australian Ocean Governance and related legislation (e.g. Coral Sea Act, The Fisheries Management Act and so on) and is developed and hosted by Geoscience Australia which is an agency of the Australian federal government. How the government backs up IMOS' Ocean Portal is also stated on the portal (http://imos.org.au/aodn.html): "IMOS is supported by the Australian Government through the National Collaborative Research Infrastructure Strategy and the Super Science Initiative." Furthermore (Moltmann et al., 2010) states that IMOS "is funded by the Australian Government through the National Collaborative Research Infrastructure Strategy (NCRIS) and the Education Investment Fund (EIF) [. . .] ". Concluded it can be stated that both initiatives are backed up with funding, development and hosting by the Australian government.

E2 – Architecture For AMSIS all that could be found is a proposed web mapping application (see figure B.3) in which OGC web map services and later feature services play an important role.

As shown in figure B.4 the service-oriented architecture of Ocean Portal is build upon three main components: a portal, a metadata catalogue and a variety of web services. The web portal offers discovery and access to the information stored in the catalogue. The catalogue has functions like the registration and management of metadata which is accessible resp. searchable through the portal or by any other client software through its OGC CSW interface. The data and services described by the catalogue are provided through the web services and can also be accessed through the portal. (Reed et al., 2007)

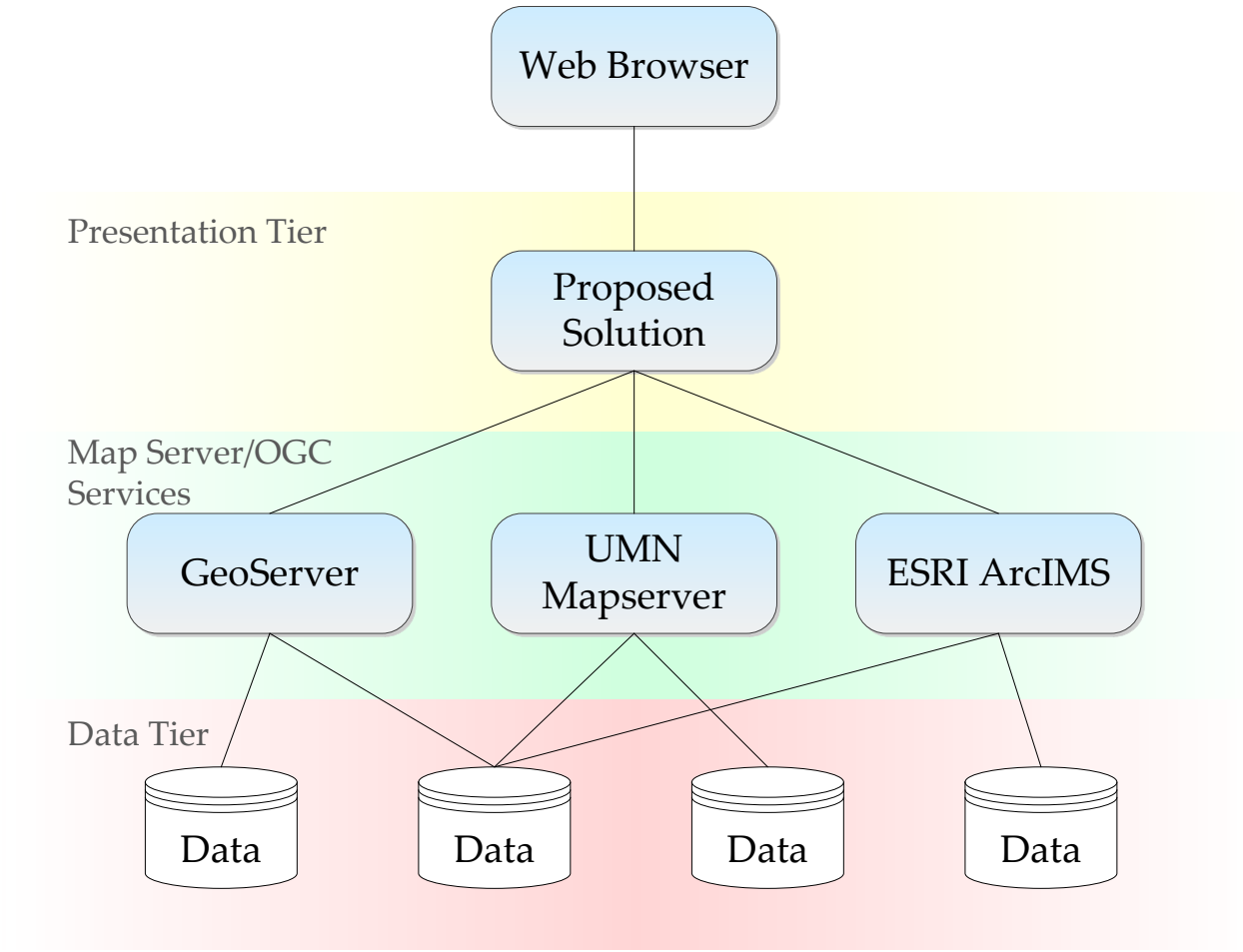


Figure B.3: *Proposed web mapping architecture for AMSIS (modified after (Nairn, 2009, p. 26))*

B.2 Canada: Marine Geospatial Data Infrastructure (MGDI), COINAtlantic and GeoPortal

A – Data The first area looks at the core datasets the Canadian approach offers and in which coordinate reference systems its data is made available.

A1 – Core datasets For the CGDI various framework datasets (Schut, 2001) are defined of which these relate to marine themes:

- *Structures:* Dataset includes significant man-made structures such as bridges, airport terminals, lighthouses, ferry terminals, ports and dams.
- *Hydrography:* Dataset includes rivers, lakes, glaciers, snowfields and coastlines. Both terrestrial and marine features are included in this layer.
- *Elevation:* Dataset includes DEM data that cover both the terrestrial and marine portions of Canada.
- *Imagery:* Dataset includes visual reference imagery.
- *International Boundaries:* The current international boundaries including marine international boundaries.
- *Ecological Units:* Two of five levels of the ecological framework (ecozones and ecoregions) are included for the marine areas.
- *Watersheds:* Dataset includes marine boundaries.

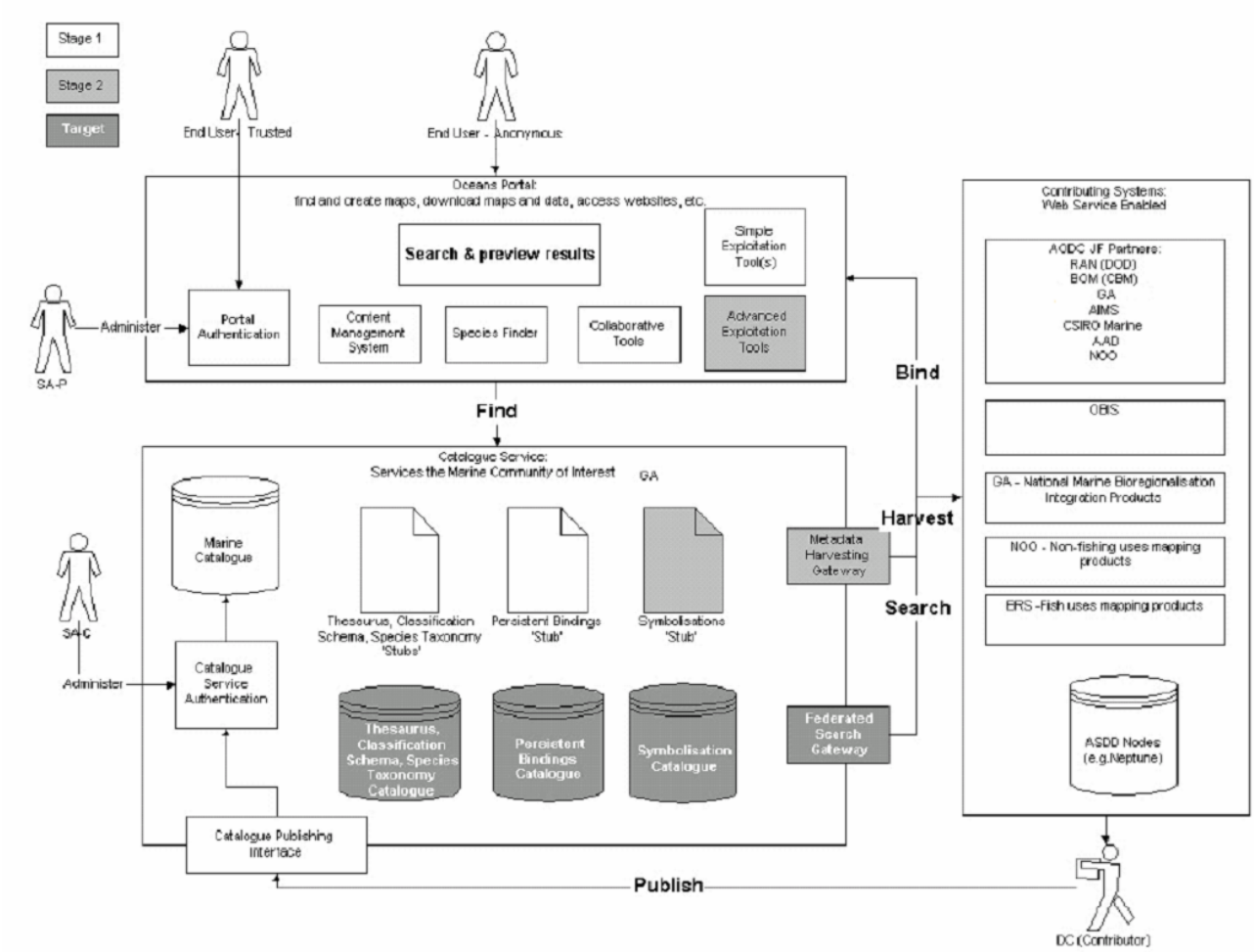


Figure B.4: *IMOS Nodes and Facilities (source: (Proctor et al., 2010, p. 13))*

As one can see bathymetry is missing (although this should be included in the CGDI's *elevation* theme) but the DFO GeoPortal is delivering the 500 meter gridded Canadian Digital Bathymetric Model (CDBM-500m) (DFO, 2008). The COINAtlantic portal provides these layers whereupon many of them are provided by the DFO GeoPortal[6]:

- International Boundaries
- NAFO (Northwest Atlantic Fisheries Organization)
- Feature names
- Power network
- Structures
- Populated places
- Railway
- Road network
- Water features
- Constructions
- Landforms
- Water saturated soils
- Hydrography
- Designated areas
- Built-up areas
- Vegetation
- Limits and Bathymetry

A2 – Degree of involvement of different agencies/institutions Nothing.

B – Metadata Area B deals with metadata – its availability, if it includes fields about data quality/accuracy and how its definition and usage is coordinated.

[6]source: http://www.marinebiodiversity.ca/coin/

B1 – Availability of Metadata/Metadata catalogue (CSW) Metadatastandardwise the CGDI (and thus the MGDI as well) moved from Z39.50 Application Profile for Geospatial Metadata (GEO) over Content Standard for Digital Geospatial Metadata (CSDGM) to today used North American Profile (NAP) of ISO 19115 (refer to indicator B on page 169). The central metadata entry point for the CGDI is the GeoConnections Discovery Portal (GDP, a component of the CGDI, http://geodiscover.cgdi.ca/). The GDP enables users to find, evaluate, access and visualize geospatial and geoscience data products and web services. For registered users it also allows publishing of data and/or services. The user is able to view found data sets resp. services – but this option is not available for all data sets – with the map viewer which is based on Ext JS (a JavaScript library). The GDP is the central entry point, so it does not only focus on marine data, but offers *Oceans and Coasts* as a category which offers 1976 data sets resp. services. A simple search for *marine* reveals just 46 data sets. The GDP is not the only metadata portal in the CGDI. The user becomes aware of this when he looks for services resp. layers from the DFO Science branch (see indicator C1 on page 171) which are not available through the GDP.

Another metadata portal is the DFO MetaData Catalogue[7] of DFO's Geoportal which serves as the central entry point for metadata discovery for DFO data. The catalogue is based on the M^{3}Cat software and it is connected to several other portals from the National Research Council (NRC) and Canadian Hydrographic Service (CHS) for example. Through an interface of the Z39.50 protocol which the catalogue uses the GeoConnections Discovery Portal (GDP) is able to harvest the metadata records which GeoPortal maintains. (DFO Canadian Science Advisory Secretariat, 2006)

The COINAtlantic initiative has primarily focused on the Canadian GeoConnections Discovery Portal (GDP) to discover the metadata (Butler et al., 2011) although COINAtlantic has its own portal for metadata discovery and geospatial data access – the COINAtlantic Search Utility (CSU) available at: http://www.marinebiodiversity.ca/coin/. Figure B.5 underlines that CSU is based on GDP.

Summing up it can be said that there is no central metadata entry point which covers all Canadian data sets covering the marine field. Regarding the availability of a CSW the CGDI offers a general CSW which is not specifically made for the marine area[8].

B2 – Data quality and accuracy In Appendix A of (Vachon et al., 2002, p. 69) which lists the mandatory attributes recommended for data collections of GeoPortal there is at least the category *Data Quality Information*. Unfortunately there does not seem to be any mandatory attributes for this category.

B3 – Coordination The CGDI's Technology Advisory Panel (TAP) is coordinating the usage of technical standards and specifications and thus evaluates and endorses these and checks if and how they should resp. can be used in the CGDI. TAP's endorsed specifications have to be used in order to be compliant with the CGDI. (GeoConnections, a) The specification endorsed for metadata is just called *Metadata for Geodata*. (GeoConnections, d) The reason for this is that there is not only one standard which can be used for metadata but there is CSDGM and concurrently NAP for ISO 19115 and there was Z39.50 before these two were introduced. (GeoConnections, b)

Z39.50 is an American National Standards Institute (ANSI)/International Organization for Standardization (ISO 23950) standard for information retrieval in interoperable way which means that catalogues are able to communicate with each other in a standardized way. The CGDI's Geodata Discovery Service used the Z39.50 Application Profile for Geospatial Metadata (GEO) from the U.S. Federal Geographic Data Committee. Z39.50 with its GEO profile defines an interface which supports search in a catalogue and which returns selected metadata records. Z39.50 is a stateful protocol which means that after the client establishes a connection with the server he sends out a search

[7] source: http://public.geoportal-geoportail.gc.ca/cat_search/public/search.jsp, retrieved: 20.12.2011

[8] It is available here: http://geodiscover.cgdi.ca/wes/serviceManagerCSW/csw?request=getCapabilities&service=CSW&version=2.0.2

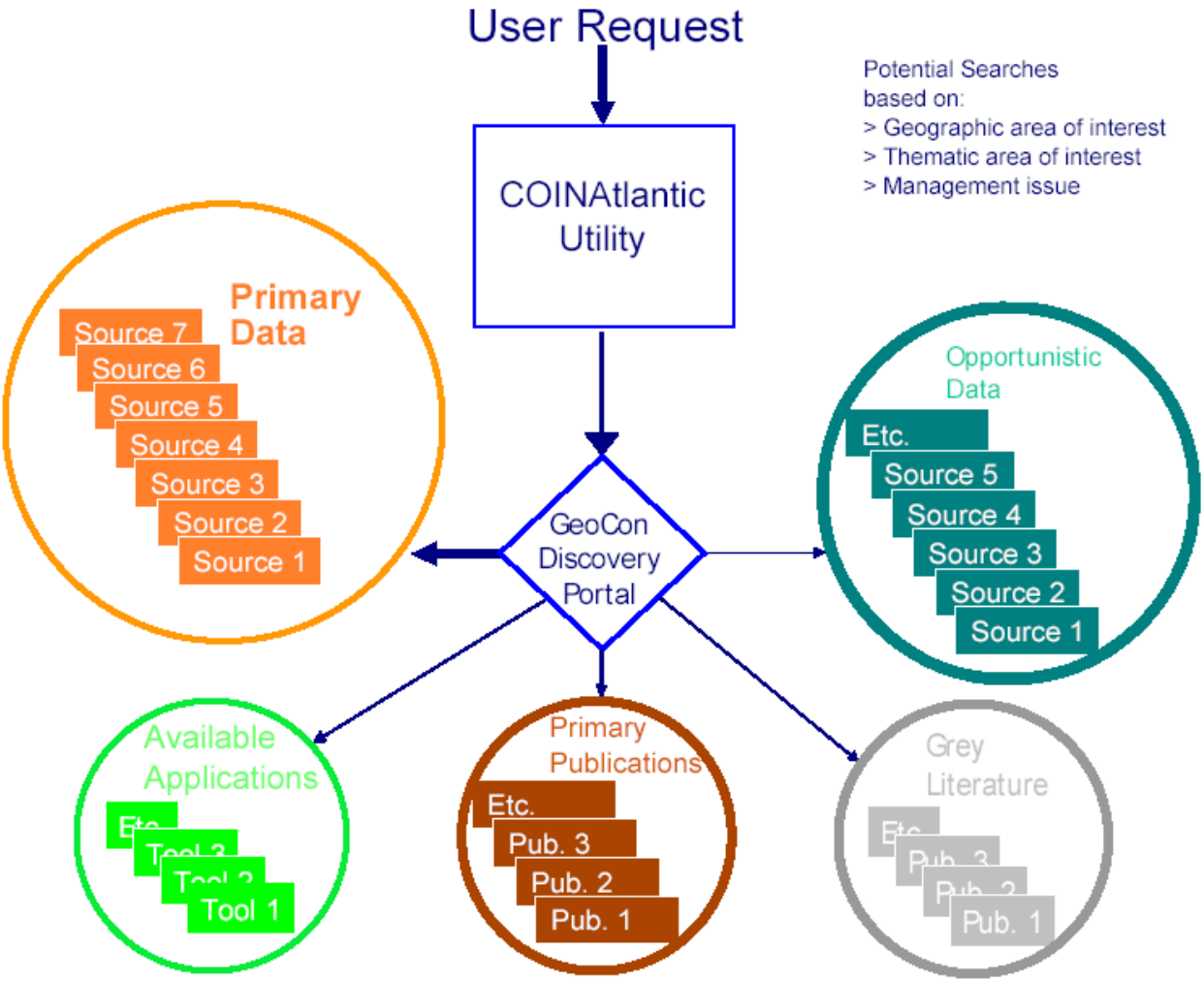


Figure B.5: *Schematic diagram of COINAtlantic search utility (source: (Butler et al., 2011, p. 11))*

request to which the server responds with search results. (GeoConnections, 2001)

The CGDI is also able to cooperate with other catalogues not using GEO but the U.S. Federal Geographic Data Committee's *Content Standard for Digital Geospatial Metadata* (CSDGM). CSDGM supports metadata profiles like the one for Shoreline Data. This profile addresses the complexities in the definition and mapping of shoreline data and data that intersects with the shoreline by providing the format and content for the description of the data sets. The profile also includes a glossary and bibliography offering a basis for understanding the shoreline and related issues. (FGDC, 1999) But CSDGM is not fitting for the metadata required for geolinked data (data which is not directly connected to a geometry/coordinates) which is important to the CGDI because important metadata fields are missing from CSDGM.

That is why the CGDI moved on to the North-American Geospatial Metadata Profile for ISO 19115 (NAP). In 1999 the ISO's Technical Committee (TC) 211 (Geographic Information/Geomatics) began the harmonization of CSDGM with other metadata standards to address new requirements for geospatial documentation. This resulted in the publication of ISO 19115 but the U.S.A. and Canada needed a broader range of capabilities to document their geospatial resources. So they started developing the NAP which adopted the 22 core elements, offered open selection of other elements, options to extend fixed domains and to increase conditionality. (FGDC, 2011)

C – Services and Interfaces Area C is concerned with services and Interfaces and looks at what Canadian marine-themed services are available, geoportal(s) and the custodianship of the data which are provided through the available services.

C1 – Availability of Services The CGDI only provides a framework for the creation of web services which means that it defines a set of abstract services to allow organizations etc. to construct components resp. systems their own services. Proposed services include but are not limited to:

- Catalogues and Registries for discovery and direct access of data
- Web Coverage Services to deliver vector datasets, raster imagery, and other types of spatial datasets
- Web Feature Servers for retrieval or editing of individual geo-spatial features
- Event Notification Services to notify applications of changes

(CGDI, 2001, p. 7)

The developed services include the CSW mentioned in indicator B1 on page 169. But there are several other web map services available in the marine area of the CGDI. The server *Atlas of Canada* (http://atlas.gc.ca/cgi-bin/atlaswms_en) provides, inter alia, the layers Ocean Background, Ocean Coastline, Bathymetry, Glaciers and Icefields, Drainage, Ice-Edge, Lakes and Reservoirs, Seaice, Transportation Network, International Boundaries, Water Areas (at different scales), Protected Areas (at different scales), Ocean Drainage Areas and Bay Names. The Canadian Aquatic Biomonitoring Network (CABIN) is also offering a web map service (http://cabin.cciw.ca/cabin_ows.asp) which offers aquatic biological monitoring data for evaluating the health of freshwater ecosystems in Canada. DFO Science branch in conjunction with Maritimes Region department provides access to diverse web map services via an ArcIMS 9.3 server (Bluefin2.dfo-mpo.gc.ca). The offered layers include (layers taken from: http://www.marinebiodiversity.ca/coin/gallery/index.php) Real time Canadian East Coast Ocean Modelling (CECOM) sea surface ice coverage model output, Gebco Isobaths, CECOM Real Time Sea Water Temperature (50M depth, surface etc.), DFO Maritimes Ocean Management Areas (OMA), Gebco gridded bathymetry and elevation, Summer groundfish trawl observations of wolffish and Real time CECOM sea current model output.

According to (Vachon et al., 2002, p. 3) DFO's GeoPortal is at least offering these services (see also figure B.6):

- A catalogue service using the M^3Cat system (see indicator B1 on page 169)
- A Cascading Web Mapping Service to allow for the inclusion of proprietary data sources such as ENC S-57 chart data, BSB raster charts and data stored in Oracle Spatial, as well as the inclusion of Web Maps coming from other OGC-compliant Web Map Servers
- A Web Feature Service that provides the capability to perform data manipulation (insert, update, delete, select) transactions on geospatial data
- A translation and download service to allow users to download data that has been identified as 'downloadable', from a user-specified geographical area, and translate it into one of many supported formats. The following data are currently available through this service:
  - Topographic data from the National Topographic Database (NTDB)
  - Hydrographic data, in the form of S-57 chart data, raster charts in BSB format, and Natural Resource Maps (NRM) in BSB format set up as common layers to be accessed by all Portal users

C2 – Performance Nothing.

C3 – Clearinghouse and geoportal Indicator C1 (on page 171) revealed that there is no central metadata catalogue on its own. The same is true for a central geoportal for data access resp. view. The CGDI has a data access point built-in the GDP (through its map viewer) while DFO's GeoPortal

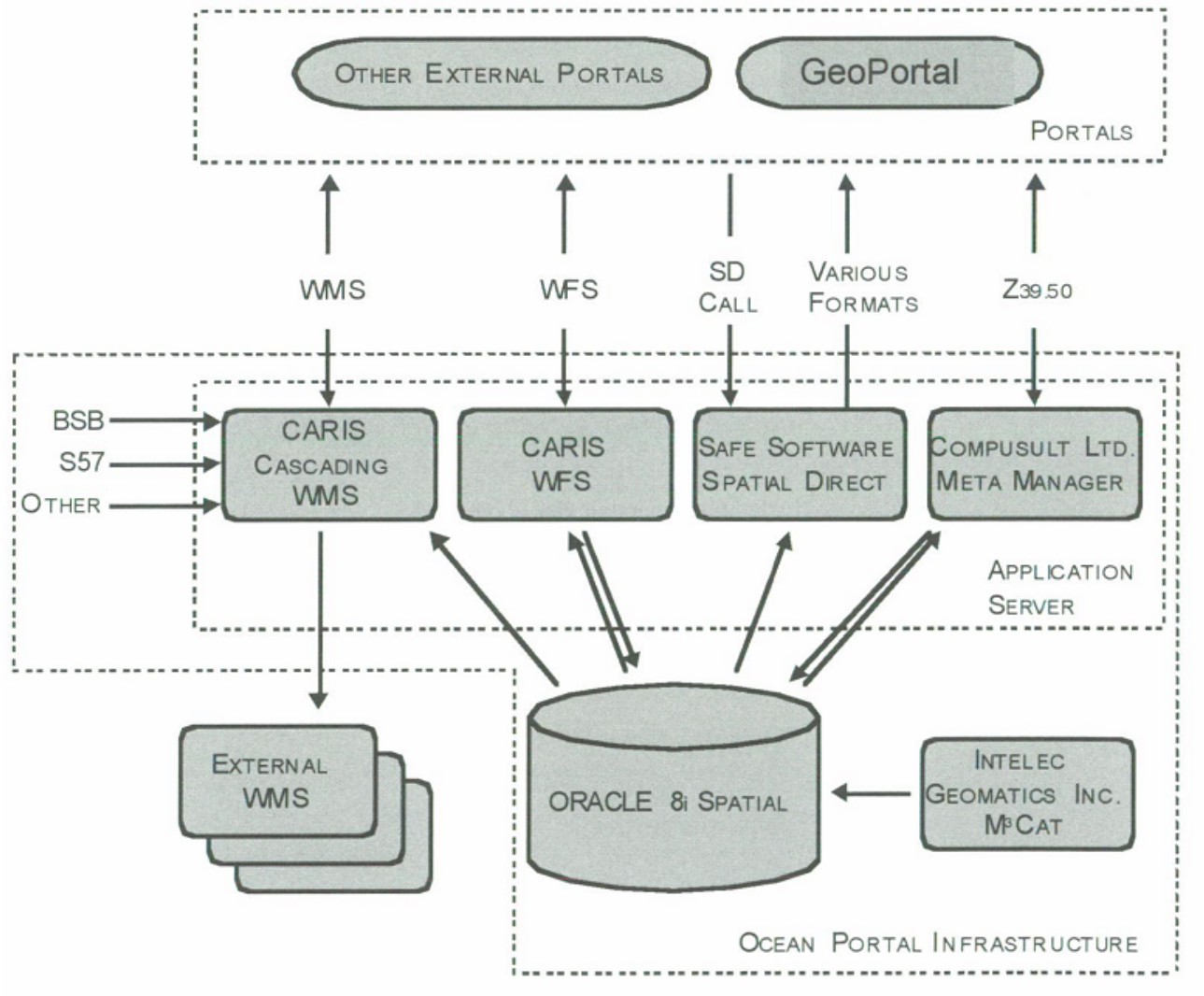


Figure B.6: *GeoPortal software components (source: (Vachon et al., 2002, p. 3))*

offers its public GeoBrowser (Department of Fisheries and Oceans Canada,). COINAtlantic is – just as the CGDI – offering its data access point together with its metadata catalogue through the same portal (http://www.marinebiodiversity.ca/coin/). All of the portals offer search functionality and a map viewer but obviously one superordinate portal which offers all the data of all three portals is lacking.

C4 – access privileges/Custodianship As already mentioned in indicator B1 on page 169 the GDP at least has two roles: unregistered user who is only able to find and access the data and the registered user who is also allowed to publish and/or modify data sets. Apart from that the MSDI (in fact the Marine Advisory Committee of the Marine Advisory Network node) identified the segments of the marine sector in which stakeholders can be found:

- Marine transportation
- Marine habitat management
- Integrated coastal zone management
- Renewable resources and biodiversity
- Non-renewable resources
- Disaster management/eme rgency response
- Sovereignty and defence
- Ocean research
- Recreation and tourism
- Freshwater resource management
- Marine engineering and works

(DFO, 2001, p. 4)

D – Standards / D1 – Interoperability As outlined in indicator B3 (on page 169) the CGDI's Technology Advisory Panel (TAP) is coordinating the usage of technical standards and specifications. The specifications endorsed by the CGDI for everything else than metadata mainly come from the Open Geospatial Consortium (OGC) because the portals are built on a web-service architecture. DFO's GeoPortal for instance at least uses the following international specifications(Vachon et al., 2002, p. 3):

- OGC Web Map Service (WMS)
- OGC Web Feature Service (WFS)
- OGC Geodata Discovery Service
- OGC Geographic Markup Language (GML)
- IHO S57

These international specifications are allowing interoperability through standard requests resp. interfaces. To achieve even more interoperability the CGDI is endorsing and/or investigating even more standards. (GeoConnections, c) lists these endorsed (must be used), suggested (should be used) or investigated (should be considered) standards divided into several categories:

- Discovering Geospatial Resources
  - Geodata Discovery Service - (Endorsed Standard)
  - Metadata for Geodata - (Endorsed Standard)
  - Catalog Services Interface - (Recommendation Paper)
- Viewing Geospatial Data
  - Web Map Service (WMS) - (Endorsed Standard)
  - Web Map Context Document - (Endorsed Standard)
  - Styled Layer Descriptor (SLD) - (Endorsed Standard)
  - Tile Mapping Service (TMS) - (Endorsed Standard)
  - Geolinking Service (GLS) - (Discussion Paper)
  - KML - (Endorsed standard)
- Accessing Geospatial Data
  - GeoRSS - (Endorsed Standard)
  - Geolinked Data Access Service (GDAS) - (Endorsed Standard)
  - Web Coverage Service (WCS) - (Endorsed Standard)
  - Web Feature Service (WFS) - (Endorsed Standard)
  - Filter Encoding - (Endorsed Standard)
  - Geographic Markup Language (GML) - (Endorsed Standard)
  - Gazetteer Service - (Discussion Paper)
- Manipulating Geospatial Data
  - Web Processing Service (WPS) - (Endorsed Standard)

E – Modelling The last area E concentrates on modelling of the Canadian approach(es) from an organizational viewpoint and reviews the existence of a government policy to back up the Canadian developments and analyses the underlying architecture.

E1 – existence of a government policy for SDI First of all the CGDI is the national spatial data infrastructure (NSDI) of Canada which means that it is implemented by the Canadian government. Furthermore the CGDI " [. . .] recognizes that governments have a responsibility to make geospatial information available [. . .]." (Labonte et al., 1998) Apart from that key participants of the twelve nodes are governmental departments resp. centres like the Department of Fisheries and Oceans (DFO) and the Canadian Centre for Marine Communications (CCMC).

All these clearly show that the government is backing up the development of the various initiatives inside the CGDI more or less directly.

E2 – Architecture The underlying scheme of the CGDI (and thus the MGDI) is the common one for a service based architecture where there are service and/or data providers, client applications and/or portals and users who work with the provided services resp. data. (CGDI, 2001)

On a more technical level the architecture is divided into three tiers which also could be considered as layers like in the OSI (Open Systems Interconnection) model because each layer provides functions, data etc. to its upper layer and uses (if applicable) functions, data etc. of the layer below (see figure B.7). The bottom layer (tier 1) deals with data management and thus we find object-relational database management systems (DBMS) there which manage the data and have the ability to operate spatial data and offer spatial functions. Furthermore the bottom layer holds data repository tools to manage metadata. Above tier 1 data access protocols and APIs are found which offer the interfaces to access the data from the bottom layer but which still belong to the bottom tier/layer. The intermediate layer (tier 2) concentrates on application servers which manage and control security, transactions, access and input from remote real-time sensors. The top layer (tier 3) offers application environments, inter alia, with client access tools to browse, query and visualize data and information, with integration tools like translation software for loading data into the data base and for converting data into other formats.

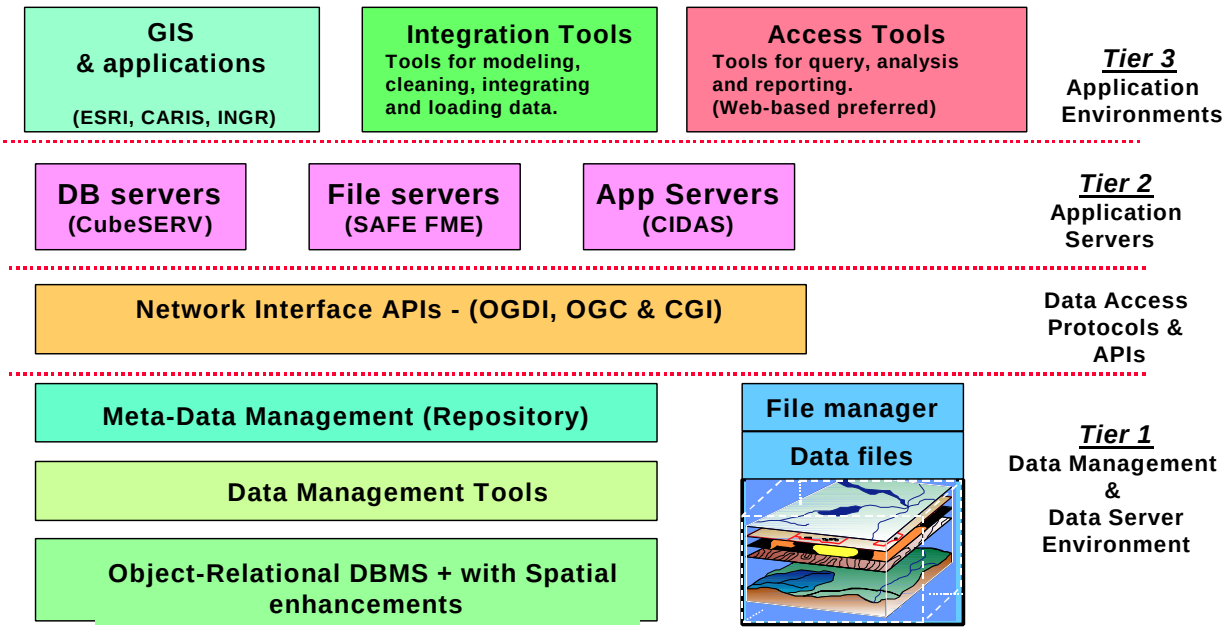


Figure B.7: *N-tier technology architecture for MGDI, including data servers (Tier 1), application servers (Tier 2) and applications (Tier 3) (source: (Gillespie et al., 2000, p. 22))*

The outlined architecture so far mainly showed the data flow. But this is not everything which is needed for a successful SDI. It also requires definition and application of:

- a common spatial data model to ensure communication between software components
- an integrated process and data modelling environment for communication between the stakeholders
- a common spatial language and data exchange format to ensure interoperability

The architecture is the foundation of the CGDI and thus forms the backbone of the MGDI.(Gillespie et al., 2000, pp. 21-22) (Sherin et al., 2009, p. 79) states that the technical architecture of COINAtlantic

will be consistent with that published by (Gillespie et al., 2000). DFO's GeoPortal is taking a similar approach like COINAtlantic and MGDI which both use the architecture suggested by the CGDI but mainly focuses on services (see figure B.8). When looking deeper into its architecture one can see that it is mainly divided into three layers: clients, portal and services whereas the services are the central component without whom the other layers could not work (see figure B.8).

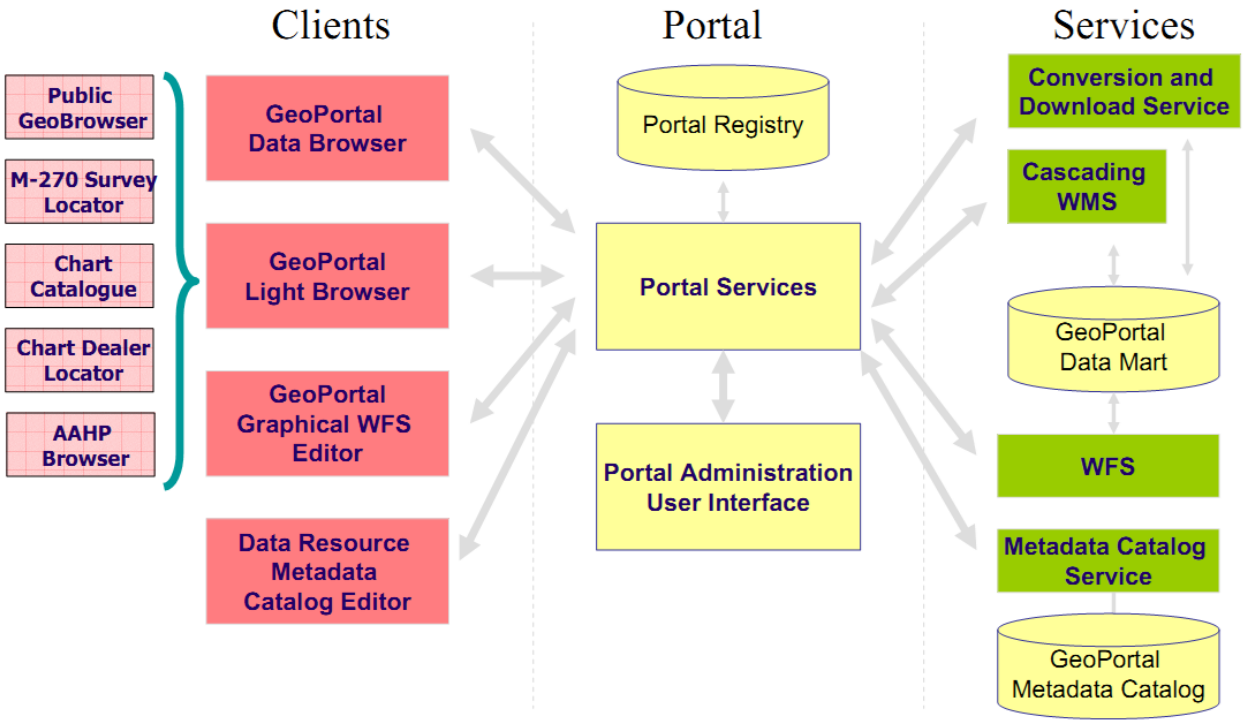


Figure B.8: *Opening the hood – GeoPortal Architecture (source: (Vachon et al., 2002, p. 12))*

B.3 USA: CMSP Registry, Marine Cadastre, Data.gov and Geoplatform

A – Data

A1 – Core datasets Overall all four examined portals (Data.Gov, CMSP, FGDC and MarineCadastre) are covering the core data sets well (see table B.1). Only the specialized cadastral data portal MarineCadastre has – with 4 out of 7 data sets – no completely positive rating.

Table B.1: *US core data sets*

Data set	Data.Gov	CMSP	FGDC	MarineCadastre
Bathymetry	+	+	+	+
Coastline and other zones	+	+	+	+
Marine cadastre	+	+	+	+
Coastal imagery	+	+	+	-
Marine navigation	+	+	+	-
Tidal heights	+	+	-	-
Nature conservation zones	+	+	+	+

A2 – Degree of involvement of different agencies/institutions The criterion for this research was how many institutions/organizations have provided data for the portal. (Lakhani et al., 2010) give

an overview of the goals and design features of the collaborations of *Data.Gov* which is reflected in table B.2. On the one hand the portals *Data.Gov* and the *MarineCadastre* have a very good cooperation with other institutions. On the other hand the portal CMSP exists only to publish its own data and the FGDC (*Geoplatform*) usually accesses the data of Data.Gov.

Table B.2: *US data sets and institutions*

	Data.Gov	**CMSP**	**FGDC**	**MarineCadastre**
Datasets (all)	73 211	-	-	-
Institutions (all)	211	-	-	-
Datasets (marine)	19 600	50	-	181
Institutions (marine)	14	1	-	11

B – Metadata

B1 – Availability of Metadata/Metadata catalogue While metadata are easily accessible on *Data.Gov* and *CMSP*, no metadata were available in the *MarineCadastre*. The evaluation of the portal FGDC was not possible for this indicator, since signing up would be required. However, it is assumed that the metadata is processed similarly to Data.Gov because Geoplatform relies on that data. (Maali et al., 2010) have also dealt with the availability of metadata on Data.Gov and come to the conclusion that for over 95 % of the data metadata is available. For the portal CMSP is has to be mentioned that the portal has no search function, but its about 50 data sets are all provided with metadata.

B2 - Coordination The portals *Data.Gov* and the *FGDC* that accesses Data.Gov are using the internationally accepted ISO standard 19115 and the original FGDC Metadata standard developed by the FGDC. The FGDC has been an important institution for the development of metadata standards since its founding in 1994 (Di et al., 2000). The use of many elements of the FGDC standards, e.g. in ISO 19115, underlines the good coordination of metadata standards in the United States.

C – Services and Interfaces

C1 – Availability of Services The available services of the different portals can be seen in table B.3.The portal of the FGDC is an exception, as this is a map portal that has been designed for display and combination of different data layers to create your own maps. Therefore this portal could not be assessed. The other portals have a relatively high number of available services, whereat for the MarineCadastre due to the lack of search function only approximate values can be given.

Table B.3: *US available services*

Service type	**Data.Gov**	**CMSP**	**FGDC**	**MarineCadastre**
WFS	306	-	-	-
WMS	610	35	-	~ 90
ESRI	815	50	-	~ 100
Gazetteer	+	-	-	-

C2 – Performance The web application Service Status Checker by the FGDC (see 3.6.3.3 on page 72) shows that the US-American approaches are very aware of performance being important.

C3 – Geoportal Table B.4 depicts the available functionalities of the geoportals. While *Data.Gov* is a central portal with geodata viewer and search function this applies only partially on the other portals. The portal equaling Data.gov is the *FGDC geospatial platform*. The *CMSP* portal and the *MarineCadastre* are missing a search functionality.

Table B.4: *US geoportals and their functionalities*

	Data.Gov	**CMSP**	**FGDC**	**MarineCadastre**
Viewer	+	+	+	+
Search	+	-	+	-

D – Standards / D1 – Interoperability The portals for which metadata are available usually the ISO 19115 standard is used , which is also used in many other countries. Apart from metadata many web services are offered as interfaces. A critical point of the U.S. geoportals is the close link with the commercial geographic information system provider ESRI. Thus, many data sets are only available only through a data viewer from the ESRI website. Especially the CMSP portal relies on a complete embedding of ESRI in their portal. The FGDC portal offers good interoperability, too. The MarineCadastre usually offers the option to download the data like the three other portals.

E – Modelling

E1 – existence of a government policy for SDI Because of the *National Ocean Policy Implementation Plan* and its National Ocean Council which proposed the focusing on Data.gov as central portal[9] a government policy is clearly existing.

E2 – Architecture Nothing.

B.4 The United Kingdom: MAGIC/CAMRA and MEDIN

A – Data

A1 – Core datasets CAMRA is mostly offering core data sets in the field of boundaries (Continental Shelf Limit (GB), Six Mile Nautical Limit (GB) etc.) while MEDIN is offering core data sets in a variety of fields like elevation and bathymetry, limits and coastlines and so on (DEFRA, 2006, pp. 13) and (Charlesworth et al., 2009, p. 6).

A2 – Degree of involvement of different agencies/institutions The Atlas (CAMRA) is the product of a project managed and collaboratively funded by the Maritime and Coastguard Agency (MCA), the Department of Environment, Food and Rural Affairs (Defra) and eleven more. MAGIC manages and distributes the data on behalf of the project partnership. According to [TMS10] on page 15 "MEDIN has 13 sponsors and many more partners that are primarily from the government sector but also include some private organisations." This includes inter alia the United Kingdom Hydrographic Office and the Department of Environment, Food and Rural Affairs (Defra). On top of that MEDIN is hosted by the British Oceanographic Data Centre (BODC). Sources:

- (WS6 Communications, 2010, p. 15)

[9] http://www.whitehouse.gov/sites/default/files/microsites/ceq/national_ocean_policy_draft_implementation_plan_01-12-12.pdf

- http://magic.defra.gov.uk/camra.html
- http://www.oceannet.org/contact_us/

A3 – Connection to EU directives As will be outlined in D1 MEDIN's Discovery Metadata Standard is compliant with INSPIRE making it easy for MEDIN to publish its metadata through a UK node and thus to INSPIRE. MEDIN's data archive centres are going to provide INSPIRE view and download service (Charlesworth, 2012, p. 15). For the Water Framework Directive (WFD) each organisation (e.g. Environment Agency [EA], Scottish Environment Protection Agency [SEPA], Northern Ireland Environment Agency [NIEA]) reports on its own but Defra plans to coordinate reporting through MEDIN (Charlesworth, 2009). There is no connection between CAMRA and EU initiatives.

B – Metadata

B1 – Availability of Metadata/Metadata catalogue CAMRA is hosted as a sub-topic within MAGIC and thus metadata in the atlas must meet the data standards adopted by MAGIC which is a metadata form. Although the forms have required and optional fields[10] according to (DEFRA, 2006, p. 21) the forms were sometimes fully completed and sometimes only a few fields were completed which would result in varying degrees of metadata availability for datasets.

MEDIN metadata has to conform to the MEDIN Metadata Discovery Standard which is compliant to other international standards (INSPIRE, ISO19115) providing the opportunity to easily exchange data between organisations. The MEDIN Metadata Discovery Standard is a marine profile of the UK GEMINI2 standard and specifies elements (mandatory, obligatory and conditional) from that standard but also mandates elements GEMINI2 does not contain and certain thesauri[11].

B2 - Coordination No information could be acquired regarding metadata coordination.

C – Services and Interfaces

C1 – Availability of Services MEDIN does not host most of the data and if so only in its Data Archive Centres which do not provide OWS. It cannot be verified if any of the data holders linked inside MEDIN are offering any OWS or not. CAMRA does not seem to offer its data as web services either.

C2 – Performance Not applicable for both initiatives.

C3 – Geoportal CAMRA/MAGIC's portal[12] is providing all the necessary tools for navigation, identification etc., a search function, a layer list and so on. MEDIN does not offer a geoportal but offers a map based search where the user can define a bounding box with a map tool and select different regions (Cotton, 2011, p. 6).

[10] http://magic.defra.gov.uk/ProjectLibrary/InfoNotes/magi0015.pdf
[11] http://www.oceannet.org/marine_data_standards/medin_disc_stnd.html
[12] http://magic.defra.gov.uk/MagicMap.aspx

D – Standards / D1 – Interoperability In very outdated information it is found that CAMRA's data model is based on the draft British Standard BS7975 which is the metadata format of the National Geospatial Data Framework (NGDF) which – according to (Swanton, 2001, pp. 5) – was supposed to *"migrate to the ISO Metadata Standard when it is ratified in late 2001/early 2002."* As already outlined in B1 MEDIN's Discovery Metadata Standard is compliant with ISO 19115 and INSPIRE and uses ISO 19139 as schema set for implementing ISO 19115. Furthermore the standard is using controlled vocabularies so that the metadata can be marine-themed. For keywords it uses the SeaDataNet vocabulary and for coordinate reference systems it uses EPSG codes. The obligation of the elements of the MEDIN discovery standard are classified into three types:

- Mandatory: Element must be filled in
- Conditional: Element must be filled in if it exists
- Optional: Element can be filled in

Sources: (Charlesworth et al., 2010, p. 6), (Charlesworth, 2012, pp. 24) and (Seeley, 2011, pp. 2)

E – Modelling

E1 – existence of a government policy for SDI MAGIC is/was funded by the Invest to Save Budget, the Department for Environment, Food and Rural Affairs (Defra) and the Office of the Deputy Prime Minister (ODPM)[13]. Many governmental organisations are involved in MEDIN, too which means that although nothing concrete could be found the government is backing up this development. Which can be underlined by the fact that a specific workshop was held so that MEDIN *"meet[s] both industry and government needs"* (WS6 Communications, 2010, p. 15).

E2 – Architecture Nothing

B.5 Germany: MDI-DE

A – Data

A1 – Core datasets

- Bathymetry (BSH Nauthis, WMS Hydrography and WMS Topography),
- Shoreline and other maritime zones (i. a. BSH Nauthis WMS Seawards Limits),
- Marine Cadastre (BSH WMS CONTIS Administration and Facilities),
- Coastal imagery (BSH WMS NAUTHIS Skin of the Earth),
- Marine navigation (BSH WMS NAUTHIS Navigational Aids and Rocks Wrecks Obstructions),
- Tidal benchmarks (WSD WMS Pegelonline and WaDaba),
- Benthic/Nature conservation habitats (BSH WMS CONTIS Administration and Facilities) and
- BSH WMS Marine Environmental Monitoring Network (MARNET).

A2 – Degree of involvement of different agencies/institutions 11 project partners and 12 associated partners covering every institution handling marine related data. The funded parties and their sub projects (SPs) in this project are (see figure B.9):

- Federal Waterways Engineering and Research Institute (Bundesanstalt für Wasserbau BAW, SP1 – "coastal engineering and coastal water protection"),

[13] http://magic.defra.gov.uk/projectsummary.htm

- German Federal Maritime and Hydrographic Agency (Bundesamt für Seeschifffahrt und Hydrographie BSH, SP2 – "protection of the marine environment"),
- German Federal Agency for Nature Conservation (Bundesamt für Naturschutz BfN, SP3 – "maritime conservation") and
- Professorship for Geodesy and Geoinformatics at Rostock University (GG, SP4 – "scientific accompanying research").

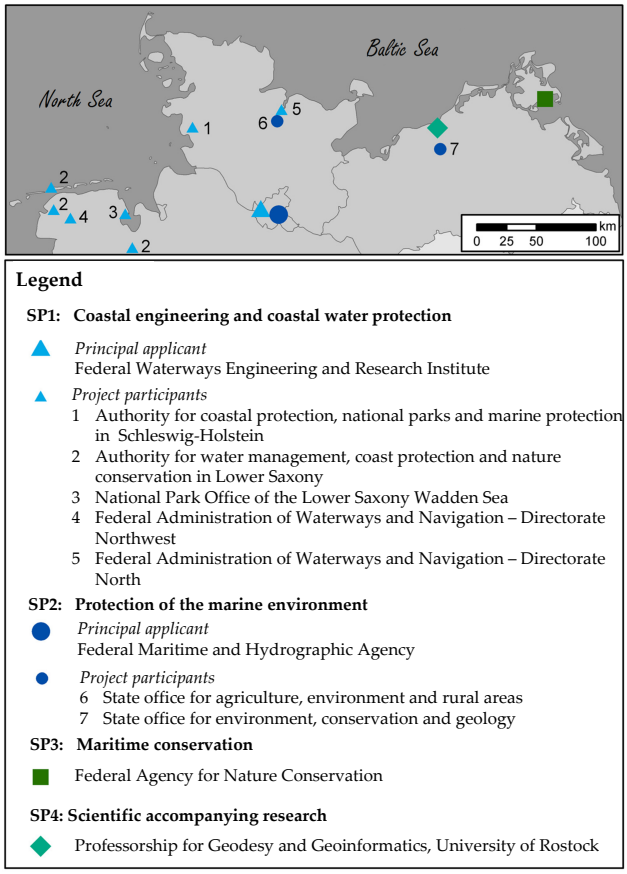


Figure B.9: *Project participants, sub projects (SP) and their locations*

A3 – Connection to EU directives Because of reporting commitments in the very near future MDI-DE focused on the Marine Strategy Framework Directive (MSFD) and a special working group (Arbeiten fÃ¼r MSRL) and another one which also produced results for the MSFD (Data harmonization and interoperability - Datenharmonisierung und Interoperabilität). Due to the rather short time span project partners in MDI-DE were only able to set up feature and map services for some of the eleven MSFD descriptors (e.g. D5 eutrophication). But developments for the Water Framework Directive (WFD) and INSPIRE were followed, too.

B – Metadata

B1 – Availability of Metadata/Metadata catalogue (CSW) Extensive metadata available for all data sets (ISO19115, ISO19119, INSPIRE and NOKIS profile) and also a central metadata catalog (terraCatalog) with CSW 2.0.2 interface.

B2 – Data quality and accuracy The availability of metadata covering accuracy of measurements for example depends on the originators of the data and because data comes from various sources the

availability is very heterogeneous. But there are several fields for data quality and accuracy in the NOKIS metadata profile for example.

B3 - Coordination The metadata is harmonized through *AG Modellierung Metadaten* and *AG Datenharmonisierung und Interoperabilität* (mostly because of MSRL requirements, this goes beyond metadata because data has to be visualized in a harmonized way ⇒ *AG Arbeiten für MSRL*).

C – Services and Interfaces

C1 – Availability of Services Up to 30 WMS services available in the portal but no WFS so far. MDI-DE will include the Gazetteer of the LKN.

C2 – Performance Evaluation see section 5.4.

C3 – Geoportal MDI-DE is more or less a central entry point for marine data in Germany. But as "more or less" indicates there are other projects resp. portals offering marine data. Most of them are specialized on research data like (PANGEA), (DeMarine), (COSYNA) and (MUDAB) but there is also the German Environmental Information Portal (Umweltportal Deutschland, PortalU) which mainly occupies itself with metadata and links to the actual data in these but it also has a map feature which offers very few marine related spatial data layers. One of these few is the census of grey seals over a period of approximately six years in the Lower Saxony Tidelands National Park (Nationalpark Niedersächsisches Wattenmeer).

D – Standards / D1 – Interoperability Standards used: ISO19115, ISO19119, INSPIRE for metadata and lots of OGC standards (CSW, WMS, WFS [Gazetteer]).

E – Modelling

E1 – existence of a government policy for SDI Yes, since it is funded by BMBF, the German Federal Ministry of Education and Research.

E2 – Architecture See 5.2.1.4.

C Theses

For the first time Germany is developing a distributed marine spatial data infrastructure (MSDI) which aims at bringing together marine related data sets collected at the federal and national level from various governmental organisations covering the North and Baltic Sea. Bringing together all information – from both thematically and sectorally-oriented areas of responsibility of the German Coastal Engineering Research Council (KFKI), marine environmental protection, marine conservation, land use planning and coastal research and responsibilities assigned to various federal and state facilities within the federal-state interaction at various levels – is a challenge. This challenge should result in a new quality both technologically and functionally as well as organizationally and legally.

However, there still is a great demand for research and development in this area. Therefore, this project of the federal and state authorities responsible for marine data – that was funded by the German Federal Ministry for Education and Research (BMBF) – was complemented by accompanying scientific-technical research and development by the Chair of Geodesy and Geoinformatics. Work for this thesis was carried out in this context with the goal to scientifically support selected aspects and essential components necessary for the general development of a MSDI.

(1) The development of a marine data infrastructure is a very complex undertaking not only because of the large number and heterogeneity of the – to be integrated – topics as well as stakeholders, but also because of the claim to not only support administrative actions. Rather, this infrastructure should support the marine sciences, the economy, policy makers and the public in order to benefit from it. This can be achieved through the use of state-of-the-art standards and a service-oriented architecture. The accompanying scientific research plays a decisive role as a neutral observer of the relevant IT developments, standardization and specification processes on ISO and OGC-level and of proving the benefits of defined components through prototypical implementations. These implementations influenced the development of MDI-DE and led to the system that was described in section 3.1 and section 5.2.

(2) This thesis proposes a framework for the evaluation of (marine) SDIs that helps to assess and compare SDIs from a rather neutral point of view (see section 4.1). The goal of the framework is to be able to learn from other existing infrastructures to find out what proved to be useful or good and where potential pitfalls are. To achieve this a set of indicators is proposed that are used to rate different systems to ensure a degree of neutrality. To prove the suitability of the resulting framework it is used to assess several national MSDIs (see section 4.2). Although strict neutrality cannot always be retained the framework proved to be suitable to achieve its goal. Thus it assists the development of MDI-DE by pointing out necessary components, designs, potential problems and so on.

(3) The foundation for the development of a MSDI is a reference model that is based on internationally and nationally accepted standards (like the Reference Model of Open Distributed Processing [RM-ODP]). This enables to specify and define the structure and composition of the infrastructure from the very beginning. Furthermore, it guarantees more transparency of the concept and of the implementation process. Lastly the reference model enhances the communication between the various actors within the infrastructure. The reference model this thesis proposes (see section 5.2) consists of several submodels (business, role, process, architecture and implementation) offering different views on the system. It is modeled with the help of the unified modeling language (UML). UML diagrams were used to model aspects of the infrastructure, i.e. scenarios, such as the evaluation of a Marine Strategy Framework Directive (MSFD) indicator. This thesis underlines that a reference model can be used to structure and specify the development of a marine SDI and model selected scenarios. However, it has to be stated that some aspects of SDI development evolve on their own nonetheless.

(4) Because of the interdisciplinary environment within an (M)SDI terms or in a broader sense vocabularies are important to clarify the meaning of terms can be used for metadata annotation, search and so on. However, despite their importance and usefulness terms are defined on various levels at several spaces and are stored in miscellaneous formats. For instance all the

existing vocabularies in the marine domain in Germany are in Excel format that cannot be used to harmonize and make the existing vocabularies available. The conversion into an advanced semantic format (e.g. Simple Knowledge Organization System [SKOS]) is needed to transfer the vocabularies to a web based thesaurus management system that enables scientists to easily exchange their knowledge, to annotate metadata consistently etc. This thesis details why SKOS was used and suggests a Free and Open Source Software (FOSS) thesauri web management tool (iQvoc, see subsection 5.6.2). It also shows that there is no tool generic enough to easily convert (Excel) vocabularies into an advanced and semantic format (see subsection 5.6.1). The existing tools are capable of executing only certain aspects of these tasks. Therefore this thesis proposes a concept for the conversion of (Excel) vocabularies into SKOS format and implements this concept (see section 5.7). With this tool it was possible to transform all three German marine-related vocabularies and import them into iQvoc. Now the terms are provided for the MDI-DE and beyond.

(5) MDI-DE is heavily influenced by reporting commitments for various European and national legislation that can be fulfilled partly using services. These services have to meet specific requirements defined by the EU and the corresponding implementation regulations regarding conformity and service quality. The services of MDI-DE were analysed using various existing tools. The usefulness of these tools to find out whether either one or both aspects needs improvement could be partly verified (see section 5.4).
   a) The results of monitoring tools help to make sure that services work as expected. Especially conformity can be evaluated rather well. However, due to varying results and the neglect of some services for various reasons it can be stated that the more tools are used the more conclusive the outcome. On the other hand service quality was not coherent when measured with different tools. Service quality proved to be not comparable and the results are not correlated to one another. This raises the question as to how the monitoring tools get, evaluate and analyse their data and if their use really makes sense, especially since each service will have to fulfil specific INSPIRE (Infrastructure for Spatial Information in the European Community) requirements in the future. Because sophistication of the tools was not subject of this thesis deficits prevail but were exposed by this thesis.
   b) Furthermore not all the monitoring tools offer visualization or a structured view of their results making them hard to interpret. This is why a visualizer web application for the results of one monitoring tool (Service Status Checker by the FGDC [Federal Geographic Data Committee, USA]) is implemented based on the JavaScript Object Notation (JSON) output of its application programming interface (API) (see section 5.5).

Aspects of the evaluation, design and implementation of (M)SDIs could be proved or enhanced through the developments and achieved results of this thesis. It has been proved that inter alia service quality still needs research and work. Furthermore developments such as JSKOSify – to convert plain thesauri into a semantic format – highlighted a gap and filled it.

D List of own publications

Korduan, Peter & Rüh, Christian: "Analyse der INSPIRE-Spezifikationen hinsichtlich der Datenverfügbarkeit für die Energiewirtschaft." In: "GeoForum MV 2012 – GIS schafft Energie: Beiträge der Geoinformationswirtschaft zur Energiewende," (edited by Ralf Bill; Guntram Flach; Ulf Klammer & Tobias Lerche), pp. 165–172. GITO Verlag, Berlin, 2012.

Lübker, Tillmann; Hübner, Peter; Rüh, Christian & Korduan, Peter: "Facilitating the Exchange of Marine Geospatial Data through the Marine Data Infrastructure for Germany (MDI-DE)." In: "Proceedings of the 26th International Cartographic Conference (ICC 2013) – From Pole to Pole," 2013.
http://www.icc2013.org/_contxt/_medien/_upload/_proceeding/268_proceeding.pdf

Rüh, Christian: "Überführung von Schlagwortlisten in SKOS." In: "Entwicklerforum Geodäsie und Geoinformationstechnik 2012: Junge Wissenschaftler forschen," pp. 89–100. Aachen : Shaker Verlag, 2012.

Rüh, Christian & Bill, Ralf: "Concepts, models and implementation of the marine spatial data infrastructure Germany (MDI-DE)." In: "ISPRS Annals - Volumes of the XXII ISPRS Congress," (edited by Mark Shortis & Marguerite Madden), volume 1-4, pp. 29–34. 2012a.

Rüh, Christian & Bill, Ralf: "A framework for evaluation of marine spatial data infrastructures - Accompanied by an international case-study -." In: "Bridging the Geographic Information Sciences: Proceedings of the 15th AGILE International Conference on Geographic Information Science, Avignon (France)," (edited by Jérôme Gensel; Didier Josselin & Danny Vandenbroucke), pp. 360–364. 2012b.

Rüh, Christian & Korduan, Peter: "Aufbau des Referenzmodells für die Marine Dateninfrastruktur Deutschlands (MDI-DE)." In: "GeoForum MV 2011 – Drehscheibe für Wirtschaft und Verwaltung," (edited by Ralf Bill; Guntram Flach; Ulf Klammer & Tobias Lerche), pp. 137–142. GITO Verlag, Berlin, 2011.

Rüh, Christian; Korduan, Peter & Bill, Ralf: "Development of the reference model for the marine spatial data infrastructure Germany (MDI-DE)." In: "Proceedings of the 25th EnviroInfo Conference," (edited by Werner Pillmann; Sven Schade & Paul Smits), pp. 419–425. Aachen: Shaker Verlag, 2011a.

Rüh, Christian; Korduan, Peter & Bill, Ralf: "A framework for evaluation of marine spatial data infrastructures to assist the development of the marine spatial data infrastructure in Germany (MDI-DE) - Accompanied by international case-studies -." In: "Proceedings of the GSDI World Conference (GSDI13)," 2012a.
http://www.gsdi.org/gsdiconf/gsdi13/papers/113.pdf

Rüh, Christian; Korduan, Peter; Bill, Ralf; Melles, Johannes; Lehfeldt, Rainer; Bauer, Michael; Hübner, Peter & Lübker, Tillmann: "Aufbau der Marinen Dateninfrastruktur Deutschland (MDI-DE)." In: *GIS-Report*, volume 2011/12: pp. 39–49, 2011b.
http://www.geobranchen.de/images/produkte/gr_11_12_ver_1.01.pdf

Rüh, Christian; Lübker, Tillmann; Binder, Kirsten; Bauer, Michael & Pramme, Matthias: "Geowebservices als Grundlage für die Erfüllung von MSRL Berichtspflichten zu Geodaten." In: "Geoinformationen für die Küstenzone Band 4," (edited by Thomas Lüllwitz Karl-Peter Traub, Jörn Kohlus), pp. 47–53. Sokrates & Freunde, Bonn, 2012b.

Rüh, Christian; Melles, Johannes & Lehfeldt, Rainer: "Marine Dateninfrastruktur Deutschland (MDI-DE)." In: *Business Geomatics*, volume 6/7, 2011c.

Rüh, Christian & Wiebensohn, Jens: "Überführung landwirtschaftlicher Schlagwortlisten in das Simple Knowledge Organization System (SKOS)." In: "Referate der 33. GIL-Jahrestagung in Potsdam 2013 – Massendatenmanagement in der Agrar- und Ernährungswirtschaft," (edited by Michael Clasen; Kurt Christian Kersebaum; Andreas Meyer-Aurich & Brigitte Theuvsen), pp. 287–290. Bonn: Köllen Verlag, 2013.

List of abbreviations

ABC	abc
ACZISC	Atlantic Coastal Zone Information Steering Committee (Canada)
AJAX	Asynchronous JavaScript and XML
AMSIS	Australian Marine Spatial Information System
ANZLIC	Australia New Zealand Land Information Council
AODN	Australian Ocean Data Network
API	Application Programming Interface
ArcIMS	Esri Arc Internet Map Server
ASDI	Australian Spatial Data Infrastructure
BAW	Federal Waterways Engineering and Research Institute (Bundesanstalt für Wasserbau)
BfG	Federal Institute of Hydrology Germany (Bundesanstalt für Gewässerkunde)
BfN	Federal Agency for Nature Conservation (Bundesamt für Naturschutz)
BODC	British Oceanographic Data Centre
BSH	Federal Maritime and Hydrographic Agency (Bundesamt für Seeschifffahrt und Hydrographie)
CABIN	Canadian Aquatic Biomonitoring Network
CAMRA	Coastal and Marine Resource Atlas (UK)
CARIS	Computer Aided Resource Information System (Marine Cartography and Geodesy Tools, Canada)
CGDI	Canadian Geospatial Data Infrastructure
CHS	Canadian Hydrographic Service
CMRC	Coastal & Marine Resources Centre at University College Cork
CMSP	Coastal and Marine Spatial Planning Data Registry (USA)
COIN	Coastal and Ocean Information Network
COINAtlantic	Coastal and Ocean Information Network for Atlantic Canada
COSYNA	Coastal Observing System for Northern and Arctic Seas
CSDGM	Content Standard for Digital Geospatial Metadata
CSDI	Coastal Spatial Data Infrastructure
CSIRO	Commonwealth Scientific and Industrial Research Organisation (Australia)
CSV	Comma-separated values
DBMS	Database management system
DCENR	Department of Communications, Energy and Natural Resources (Ireland)
Defra	Department for Environment, Food and Rural Affairs (UK)
DEM	Digital elevation model
DERM	Digital Earth Reference Model
DFO	(Department of) Fisheries and Oceans Canada
DOM	Document Object Model

EAF	Ecosystem Approach to Fisheries
EBF	Ecosystem Based Management
EEZ	Exclusive economic zone
ENC	Electronic navigational chart
EPA	Environmental Protection Agency
EPSG	European Petroleum Survey Group
EU	European Union
FFH	Fauna and flora directive (Fauna-Flora-Habitat-Richtlinie)
FGDC	Federal Geographic Data Committee
FOSS	Free and Open-Source Software
GCMD	Global Change Master Directory
GDI-DE	Spatial data infrastructure for Germany
GDP	GeoConnections Discovery Portal (Canada)
GEMET	GEneral Multilingual Environmental Thesaurus
GeoConnections	A nation-wide program of the federal Department of Natural Resources, Canada
GMES	Global Monitoring for Environment and Security (now Copernicus)
GML	Geography Markup Language
GOOS	Global Ocean Observing System
GUI	Graphical user interface
HELCOM	Baltic Marine Environment Protection Commission (Helsinki Commission)
HydroML	Hydrologic Markup Language
ICM	Integrated Coastal Management
ICOM	Integrated Coastal and Ocean Management
ICZM	Integrated Coastal Zone Management
IHO	International Hydrographic Organisation
IMOS	Integrated Marine Observing System
INSPIRE	Infrastructure for Spatial Information in the European Community
IOC	Intergovernmental Oceanographic Commission
IODE	International Oceanographic Data and Information Exchange
IOOS	Integrated Ocean Observing System
IRLOGI	Irish Organisation for Geographic Information
ISO	International Organization for Standardization
JDOM	*(not an acronym)* Java API for XML documents
JSON	JavaScript Object Notation
JSONP	JSON with Padding
LANIS	Landscape and Nature Protection Information System (Landschafts- und Naturschutzinformationssystem)
LHM	LANIS Habitat Mare
LKN	Agency for Coastal Defence, National Park and Marine Conservation (Landesbetrieb Küstenschutz, Nationalpark und Meeresschutz)
LUNG	Agency for the Environment, Nature Conservation and Geology (Landesamt für Umwelt, Naturschutz und Geologie)

MAGIC	Multi-Agency Geographic Information for the Countryside (UK)
MDI-DE	German marine spatial data infrastructure (Marine Dateninfrastuktur Deutschland)
MEDIN	Marine Environmental Data and Information Network (UK)
MGDI	Marine Geospatial Data Infrastructure (Canada)
MMI	Marine Metadata Interoperability Project
MSDI	Marine Spatial Data Infrastructure
MSFD	Marine Strategy Framework Directive
Natura2000	Ecological network of protected areas (EU)
NCRIS	National Collaborative Research Infrastructure Strategy (Australia)
NOAA	National Ocean and Atmospheric Administration (USA)
NOKIS	North-Baltic-Sea-Coastal-Information-System (Nord-Ostsee-Küsten-Informations-System)
NRC	National Research Council Canada
OASIS	Organization for the Advancement of Structured Information Standards
OGC	Open Geospatial Consortium
OSM	OpenStreetMap
OSPAR	Convention for the Protection of the Marine Environment of the North-East Atlantic (Oslo-Paris-Convention)
OWL	Web Ontology Language
OWS	OGC Web Service
PPMCC	Pearson product-moment correlation coefficient
QoS	Quality of Service
RDF	Resource Description Framework
RM-ODP	Reference Model of Open Distributed Processing
SDI	Spatial Data Infrastructure
SEIS	Shared Environmental Information System
SKOS	Simple Knowledge Organisation System
SOA	Service-oriented architecture
SQL	Structured Query Language
SSC	(FGDC) Service Status Checker
UBA	Federal Environmental Agency (Umweltbundesamt)
UML	Unified Modeling Language
UMTHES	Environmental-(Umwelt-)Thesaurus by the UBA
UNCLOS	United Nations Convention on the Law of the Sea
VS-RL	Birds Directive (Vogelschutzrichtlinie)
W3C	World Wide Web Consortium
WFD	Water Framework Directive
WFS	Web Feature Service
WISE	Water Information System for Europe
WMS	Web Map Service
WPS	Web Processing Service
WRON	Water Resources Observation Network (Australia)

WSDL	Web Services Description Language
XHR	XMLHttpRequest
XML	Extensible Markup Language
XSLT	Extensible Stylesheet Language Transformations

List of Tables

List of Figures

List of Listings

Index

Bibliography

Aime, Andrea & McKenna, Jeff: "WMS Performance Shootout." Presentation at FOSS4G 2009, 2009. http://download.osgeo.org/osgeo/foss4g/2009/SPREP/3Fri/Parkside%20Auditorium/1500/fri%20paud%201500%20mckenna-aime.pdf

Altmaier, Angela; Fitzke, Jens; Greve, Klaus & Petersen, Kerstin: "Geodateninfrastruktur Nordrhein-Westfalen – Referenzmodell." Technical report, Initiative GDI NRW, 2003. http://www.bezreg-koeln.nrw.de/brk_internet/organisation/abteilung07/dezernat_74/geodateninfrastruktur/archiv/dokumente_referenzmodell_3_1.pdf

Anthony, Michelle & Nebert, Doug: "Monitoring the Performance and Reliability of Geospatial Web Services – Service Status Checker (SSC) System Overview." Presentation at Global Geospatial Conference (GSDI13), 2012. http://www.gsdi.org/gsdiconf/gsdi13/slides/1.4c.pdf

ANZLIC Metadata Working Group: "ANZLIC Metadata Guidelines: Core metadata elements for georaphic data in Australia and New Zealand." Technical Report Version 2, ANZLIC, 2001. http://www.ga.gov.au/image_cache/GA9364.pdf

Bandholtz, Thomas & Glaser, Robert: "iQvoc – Service Documentation (Preview)." Technical report, innoQ, 2012.

Bandholtz, Thomas; Schulte-Coerne, Till; Glaser, Robert; Fock, Joachim & Keller, Tim: "iQvoc – Open Source SKOS(XL) Maintenance and Publishing Tool." In: "Proceedings of Sixth Workshop on Scripting and Development for the Semantic Web at ESWC 2010, Heraklion, Greece.", (edited by Gunnar Aastrand Grimnes; Sören Auer & Gregory Todd Williams), volume 699 of *CEUR Workshop Proceedings ISSN 1613-0073*. 2010. http://CEUR-WS.org/Vol-699/Paper2.pdf

Bartelme, Norbert: *Geoinformatik - Modelle, Strukturen, Funktionen*. Springer, Berlin Heidelberg, 2005.

Bartlett, Darius; Longhorn, Roger & Garriga, Maria Carmen: "Marine and Coastal Data Infrastructures: a missing piece in the SDI puzzle?" In: "Proceedings of the 7th Global Spatial Data Infrastructure conference," 2004.

Bernard, Lars: "Experiences from an implementation Testbed to set up a national SDI." In: "5th AGILE Conference on Geographic Information Science 2002," (edited by Maurici Ruiz; Michael Gould & Jerònia Ramon), pp. 315–321. 2002. http://plone.itc.nl/agile_old/Conference/mallorca2002/proceedings/dia26/Session_7/s7_Bernard.pdf

Bernard, Lars; Crompvoets, Joep & Fitzke, Jens: "Geodateninfrastrukturen - ein Überblick." In: "Geodateninfrastrukturen - Grundlagen und Anwendungen," (edited by Lars Bernard; Jens Fitzke & Roland M. Wagner), pp. 3–8. Wichmann, 2005. http://www.vde-verlag.de/buecher/537395/geodateninfrastruktur.html

Berners-Lee, Tim: "Linked Data." Personal note filed under "Design Issues" on W3C page, 2003. Accessed: January 31, 2014. http://www.w3.org/DesignIssues/LinkedData.html

Bill, Ralf: *Grundlagen der Geo-Informationssysteme*. Wichmann, Berlin, 2010.

Böhme, Sven & Hogrebe, Daniela: "Anwender-Handbuch GDI-DE Testsuite." user manual, Coordination Office SDI, 2012. http://testsuite.gdi-de.org/gdi/files/Anwender-Handbuch_GDI-DE-Testsuite.pdf

Bray, Tim; Paoli, Jean; Sperberg-McQueen, C.M.; Maler, Eve & Yergeau, François: "Extensible Markup Language (XML) 1.0." Technical report, W3C, 2008. http://www.w3.org/TR/xml/

Brox, Christoph; Bishr, Yaser; Senkler, Kristian; Zens, Katharina & Kuhn, Werner: "Toward a geospatial data infrastructure for Northrhine-Westphalia." In: *Computers, Environment and Urban Systems*, volume 26: pp. 19–37, 2002.
http://gdi-nrw.uni-muenster.de/publications/Brox_Geospatial.pdf

Buglione, Luigi & Abran, Alain: "Assessment of Measurement Indicators in Software Process Improvement Frameworks." In: "13th International Workshop on Software Measurement – IWSM 2003, Montréal (Canada)," pp. 287–310. 2003.
http://www.gelog.etsmtl.ca/publications/pdf/783.pdf

Busznyák, János; van Diepen, Frans; Rauthe, Norma; Escobedo, Carolina; Angileri, Vincenzo; Natale, Fabrizio & López, Angel: "Data Specification on Agricultural and aquaculture facilities Version 3.0.0 rc3." 2013.
http://inspire.jrc.ec.europa.eu/documents/Data_Specifications/INSPIRE_DataSpecification_AF_v3.0rc3.pdf

Butler, Michael J. A.; Boudreau, Paul R.; LeBlanc, Claudette & Baldwin, Kim: *Towards Marine Ecosystem-based Management in the Wider Caribbean*, chapter Spatial Data infrastructures in Support of Ecosystem Based Management and Ecosystem Approach to Fisheries in the Caribbean. Amsterdam University Press, 2011.
http://marineaffairsprogram.dal.ca/Files/Butler_Spatial_Data_infrastructures.doc

Castano, Silvana; Antonellis, Valeria De; Fugini, Maria Grazia & Pernici, Barbara: "Conceptual schema analysis: techniques and applications." In: *ACM Transactions on Database Systems*, volume 23 (3): pp. 286–333, 1998. ISSN 0362-5915.
http://doi.acm.org/10.1145/293910.293150

CGDI: "Canadian Geospatial Data Infrastructure - Target Vision." Technical report, CGDI Architecture Working Group, 2001.

Charlesworth, Mark: "Analysis of reporting of marine data/information to Europe (Draft)." Technical report, BODC, 2009.
http://www.oceannet.org/library/work_stream_documents/documents/uk_reporting_to_europe.doc

Charlesworth, Mark: "MEDIN Introduction and Outline of Discovery Metadata Standards." In: "Marine Data Standards Workshop 28th March 2012 - Edinburgh," 2012.
http://www.oceannet.org/marine_data_standards/workshop_presentations.html

Charlesworth, Mark; Allen, Terry; Cotton, David; Evans, Gaynor & Rickards, Lesley: "The UK Marine Environmental Data and Information Network: Working to Deliver Improved Access To and Stewardship of Marine Data and Information." In: "Proceedings of the GSDI 11 World Conference - Spatial Data Infrastructure Convergence: Building SDI Bridges to address Global Challenges," 2009.
http://www.gsdi.org/gsdiconf/gsdi11/slides/fri/6.3b.pdf

Charlesworth, Mark; Lowry, Roy; Freeman, Hannah; Rapaport, James & Seely, Becky: "The UK Marine Environmental Data and Information Network – MEDIN." In: "Proceedings of workshop on "Application of Standards for GI", Southampton on 24th May 2010," 2010.
http://www.isotc211.org/WorkshopSouthampton/Workshop_presentations/WS_02_Charlesworth.pdf

Christl, Arnulf: "About Data Lifetime." Entry on personal blog, 2013. Accessed: January 31, 2014.
http://arnulf.us/sevendipity/archives/59-About-Data-Lifetime.html

Cibulka, Dusan: "Performance Testing of Web Map Services tn three Dimensions – X, Y, Scale." In: *Slovak Journal of Civil Engineering*, volume XXI (1): pp. 31–36, 2013.
http://www.degruyter.com/view/j/sjce.2013.21.issue-1/sjce-2013-0005/sjce-2013-0005.xml

con terra: "serviceMonitor (Version 3.0) – Qualitätsmanagement in Geodateninfrastrukturen." Flyer, 2013.
http://www.conterra.de/de/service/download/sdi-suite/Flyer_serviceMonitor_3.0_DE.pdf

Cotton, David: "MEDIN Resources Linking to the UK Location Programme / data.gov and INSPIRE." Technical report, MEDIN Core Team, 2011.
http://www.oceannet.org/library/work_stream_documents/documents/medin_and_uklp_datagovukv1.doc

Craglia, Massimo: "INSPIRE Metadata Implementing Rules: Technical Guidelines based on EN ISO 19115 and EN ISO 19119." 2010a.
http://inspire.jrc.ec.europa.eu/documents/Metadata/INSPIRE_MD_IR_and_ISO_v1_2_20100616.pdf

Craglia, Max: "Building INSPIRE: The Spatial Data Infrastructure for Europe." In: *ArcNews*, volume Spring: pp. 5–7, 2010b.
http://www.esri.com/news/arcnews/spring10articles/building-inspire.html

Craglia, Max; de Bie, Kees; Jackson, Davina; Pesaresi, Martino; Remetey-Fülöpp, Gábor; Wang, Changlin; Annoni, Alessandro; Bian, Ling; Campbell, Fred; Ehlers, Manfred; van Genderen, John; Goodchild, Michael; Guo, Huadong; Lewis, Anthony; Simpson, Richard; Skidmore, Andrew & Woodgate, Peter: "Digital Earth 2020: towards the vision for the next decade." In: *International Journal of Digital Earth*, volume 5 (1): pp. 4–21, 2012.
http://www.tandfonline.com/doi/abs/10.1080/17538947.2011.638500

Daffner, Franz; Grellet, Sylvain; Rudolf, Heino; Abenius, Johan; Alcorn, Trevor; Bertrand, Nicolas; Lutz-Holzhauer, Christiane; Pastorek, Peter; Schleidt, Katharina; Sarretta, Alessandro; Reznik, Tomas; Lutz, Michael & Cetl, Vlado: "Data Specification on Environmental Monitoring Facilities Version 3.0.0 rc3." 2013.
http://inspire.jrc.ec.europa.eu/documents/Data_Specifications/INSPIRE_DataSpecification_EF_v3.0rc3.pdf

Danko, David M.: "Metadata and Interoperability, Geospatial." In: "Encyclopedia of GIS," pp. 656–660. Springer US, 2008. ISBN 978-0-387-30858-6.
http://dx.doi.org/10.1007/978-0-387-35973-1_459

Daws, Sharon: "Tutorial: Flot – How to Create Bar Charts." www.pikemere.co.uk/blog/, 2012. Accessed: January 31, 2014.
http://www.pikemere.co.uk/blog/tutorial-flot-how-to-create-bar-charts/

de la Beaujardiere, Jeff: "OpenGIS Web Map Server Implementation Specification 1.3.0." 2006.

de La Beaujardière, Jeff; Beegle-Krause, C. J.; Bermudez, Luis; Hankin, Steven; Hazard, Lisa; Howlett, Eoin; Le, Steven; Proctor, Roger; Signell, Richard P.; Snowden, Derrick & Thomas, Julie: "Ocean and Coastal Data Management." In: "Proceedings of OceanObs'09: Sustained Ocean Observations and Information for Society," 2009.
http://www.oceanobs09.net/proceedings/cwp/cwp22/index.php

DEFRA: "Coastal and Marine Resource Atlas - Final Report." Technical report, Department of Environment, Food, and Rural Affairs (DEFRA), 2006.
http://magic.defra.gov.uk/Me1302_4030_FRP.pdf

Department of Fisheries and Oceans Canada: "DFO GeoPortal." Accessed: June 6, 2013.
http://public.geoportal-geoportail.gc.ca/dfoGeoPortal/

DFO: "Marine User Requirements for Geospatial Data - Summary 2001." Technical report, Department of Fisheries and Oceans Canada, 2001.
www.geoconnections.org/publications/reports/marine/Marine_User_Requirements_E.pdf

DFO: "Canadian Digital Bathymetric Model Gridded at 500 meters - Product Specifications (Edition 4.0)." Technical report, Department of Fisheries and Oceans Canada, 2008.
http://www.charts.gc.ca/data-gestion/bathy/specifications-eng.asp

DFO Canadian Science Advisory Secretariat (Editor): *Using Metadata Standards to Achieve Data Interoperability: Proceedings of a Technical Symposium and Workshop*. DFO, 2006.

Di, Liping; Schlesinger, Barry M. & Kobler, Ben: "US FGDC Content Standard for Digital Geospatial Metadata: Extensions for Remote Sensing Metadata." In: "International archives of photogrammetry and remote sensing," volume XXXIII, pp. 78–81. 2000.
http://www.isprs.org/proceedings/XXXIII/congress/part1/78_XXXIII-part1.pdf

Dreesmann, Michael; Häner, Rainer; Köhler, Petra; Lochter, Frank; Stiehler, Siegmund; Strehmel, Ralf; Wächter, Joachim & Walther, Axel: "Geodaten-Infrastruktur Brandenburg (GIB) – Referenzmodell, Version 1.0." Technical report, Geodaten-Infrastruktur Brandenburg (GIB), 2004.
http://gdi.berlin-brandenburg.de/papers/GIB-Referenzmodell-1.0.pdf

Drerup, Klaus: "Setup and test of a WFS for NatureSDI+ according to INSPIRE." Bachelor Thesis, 2010.
http://hig.diva-portal.org/smash/record.jsf?pid=diva2:354500

DuCharme, Bob: "What SKOS-XL adds to SKOS." bobdc.blog, 2011. Accessed: January 31, 2014.
http://www.snee.com/bobdc.blog/2011/02/what-skos-xl-adds-to-skos.html

Dwyer, Edward; Köpke, Kathrin; Cummins, Valerie; O'Dea, Elizabeth & Dunne, Declan: *Coastal Informatics: Web Atlas Design and Implementation*, chapter 7 - Ireland, pp. 105–130. 2011.

Dwyer, Edward; O'Dea, Liz & Cummins, Valerie: "The Marine Irish Digital Atlas: A Web Portal to Coastal and Marine Data in Ireland." In: "Proceedings from CoastGIS 2003," 2003.

Dwyer, Edward & Wright, Dawn: "Report of International Coastal Atlas Network Workshop 3 on Federated Coastal Atlases: Building on the Interoperable Approach." Technical report, European Environment Agency, Copenhagen, Denmark, 2008.
http://ican.science.oregonstate.edu/ican3_final_rpt

European Commission: "INSPIRE Geoportal Metadata Validator." Technical report, 2013.
http://inspire-geoportal.ec.europa.eu/validator2/

Evans, John D.: "Digital Earth Reference Model." Technical report, NASA Digital Earth Office, 2001.
http://www.cartome.org/draft-derm.htm

Fabian, Peter: "Environmental spatial data infrastructure in Slovakia." In: "Proceedings of the 9th EC-GI & GIS Workshop (ECGIS), A Coruña, Spain," 2003.

FGDC: "Shoreline Metadata and Glossary in Support of a National Shoreline Data Content Standard." Technical report, Federal Geographic Data Committee (FGDC), 1999.
http://www.fgdc.gov/standards/projects/FGDC-standards-projects/metadata/shoreline-metadata/shore_pr.pdf

FGDC: "Preparing for International Metadata." Technical report, Federal Geographic Data Committee (FGDC), 2011.
http://www.fgdc.gov/metadata/documents/preparing-for-international-metadata-guidance.pdf

Flasbarth, Jochen: "Water Framework Directive: the way towards healthy waters." Flyer, Federal Environment Agency (Umweltbundesamt UBA), 2011.
http://www.umweltdaten.de/publikationen/fpdf-l/4194.pdf

Gao, Sheng; Coleman, David; MacLachlan, Craig & Plunkett, Gordon: "CGDI in action: exploring quality of service." In: "ISPRS Archives – Volume XXXVIII - Part 1, 2010," 2010.
http://www.isprs.org/proceedings/XXXVIII/part1/02/02_01_Paper_79.pdf

GDI-Sachsen: "Referenzmodell der Geodateninfrastruktur Sachsen." Technical report, gdi.initiative.sachsen, 2009. Version: 1.0.
http://www.gdi.sachsen.de/inhalt/download/doku/Architekturkonzept-M-GDISN_091028.pdf

GeoConnections: "Explanation of Service Specification Status Levels." a. Accessed: June 6, 2013.
http://www.geoconnections.org/en/communities/developers/standards/statusExplanation

GeoConnections: "MetaData for Geodata." b. Accessed: June 6, 2013. http://geoconnections.org/en/communities/developers/standards/metadata_for_geodata

GeoConnections: "Standards." c. Accessed: June 6, 2013. http://www.geoconnections.org/en/communities/developers/standards

GeoConnections: "Standards and specifications ... more." d. Accessed: June 6, 2013. http://www.geoconnections.org/en/communities/developers/standards/indexToSpecs

GeoConnections: "General description of the CGDI Geodata Discovery Service - Z39.50." Technical report, GeoConnections Technology Advisory Panel, 2001. geoconnections.org/architecture/technical/specifications/geodata_registry/geodata_discovery_service_z39_50.pdf

Gillespie, Randy; Butler, Michael; Anderson, Neil; Kucera, Henry & LeBlanc, Claudette: "MGDI: An Information Infrastructure to Support Integrated Coastal Management in Canada." In: *GeoCoast*, volume 1 (1): pp. 15–24, 2000.

Gore, Albert Arnold: "The Digital Earth: understanding our planet in the 21st century." In: *Photogrammetric Engineering and Remote Sensing*, volume 65 (5): p. 528, 1999.

Grohmann, Jan & Stahl, Roland: "Betriebsmodell GDI-DE." Technical report, KSt. GDI-DE, 2010. https://network.gdi-de.org/download/attachments/1442143/Projektskizze_Betriebsmodell_GDI-DE_V_01.01.pdf

Groot, Richard: "Spatial data infrastructure (SDI) for sustainable land management." In: *ITC Journal*, volume 3-4: pp. 287–294, 1997.

Groot, Richard & McLaughlin, John Douglas (Editors): *Geospatial Data Infrastructure: Concepts, Cases, and Good Practice*. Spatial Information Systems Series. Oxford University Press, 2000. ISBN 9780198233817. http://books.google.de/books?id=-IB5QgAACAAJ

Gruber, Tom: "Ontology." In: "Encyclopedia of Database Systems," (edited by Ling Liu & M. Tamer Özsu), pp. 1963–1965. Springer US, 2009. ISBN 978-0-387-39940-9. http://tomgruber.org/writing/ontology-definition-2007.htm

Guptill, Stephen C.: "Metadata and data catalogues." In: *Geographical Information Systems. Wiley, Chichester*, pp. 677–692, 1999. http://www.geos.ed.ac.uk/~gisteac/gis_book_abridged/files/ch49.pdf

Harold, Elliotte Rusty: *Processing XML with Java: A Guide to SAX, DOM, JDOM, JAXP, and TrAX*. Reading, Massachusetts: Addison Wesley, 2002. ISBN 0201771861.

Harrison, Matthew; Thomas, Florian; Barredo, José I.; Bojilov, Venco; Castella, Raquel Canet; Cerba, Otakar; Exadaktylos, George; Isidro, Miguel Llorente; Pfeiffer, Manuela & Tomas, Robert: "Data Specification on Natural Risk Zones Version 3.0.0 rc3." 2013. http://inspire.jrc.ec.europa.eu/documents/Data_Specifications/INSPIRE_DataSpecification_NZ_v3.0rc3.pdf

Heiler, Sandra: "Semantic interoperability." In: *ACM Comput. Surv.*, volume 27 (2): pp. 271–273, 1995. ISSN 0360-0300. http://doi.acm.org/10.1145/210376.210392

Hinterlang, Dirk; Friis-Christensen, Anders; Bruns, Peteris; Carlisle, Margaret; Claus, Simon; Jongman, Robert; Kliment, Tomas; Sharry, Brian Mac; Maxim, Lurie; May, Rudolf; Peterseil, Johannes; Roscher, Sabine; Schmidt, Anne; Ssymank, Axel; Tirry, Diederik; Valland, Nils; Vischer-Leopold, Mareike; Wilkinson, Steve; Tuchyna, Martin & Abramic, Andrej: "Data Specification on Habitats and Biotopes Version 3.0.0 rc3." 2013. http://inspire.jrc.ec.europa.eu/documents/Data_Specifications/INSPIRE_DataSpecification_HB_v3.0rc3.pdf

Hjelmager, Jan; Moellering, Harold; Cooper, Antony; Delgado, Tatiana; Rajabifard, Abbas; Rapant, Petr; Danko, David; Huet, Michel; Laurent, Dominique; Aalders, Henri; Iwaniak, Adam; Abad, Paloma; Düren, Ulrich & Martynenko, Alexander: "An initial formal model for spatial data infrastructures."

In: *International Journal of Geographical Information Science*, volume 22 (11&12): pp. 1295–1309, 2008.
http://www.csdila.unimelb.edu.au/publication/journals/AN%20INITIAL%20FORMAL%20MODEL%20FOR%20SDI.pdf

Hogrebe, Daniela: "Testing conformity of services and datasets within SDI Germany (GDI-DE Testsuite)." Project outline, Coordination Office SDI, 2011.
https://wiki.gdi-de.org/pages/viewpage.action?pageId=5177418

Hogrebe, Daniela: "GDI-DE Testsuite. Improving interoperability." Presentation at the INSPIRE conference 2012, 2012.
http://inspire.jrc.ec.europa.eu/events/conferences/inspire_2012/presentations/90.pdf

Horák, Jirí; Ardielli, Jirí & Horáková, Bronislava: "Testing of Web Map Services." In: "Proceedings of the GSDI 11 World Conference - Spatial Data Infrastructure Convergence: Building SDI Bridges to address Global Challenges," 2009.
http://www.gsdi.org/gsdi11/papers/pdf/330.pdf

Hower, Rick: "Web Site Test Tools and Site Management Tools – Load and Performance Test Tools." Website, 2013. Accessed: January 31, 2014.
http://www.softwareqatest.com/qatweb1.html#LOAD

Huynh, David: "Google Refine Tutorial." Presentation at CAR 2011 conference, 2011.
http://davidhuynh.net/spaces/nicar2011/tutorial.pdf

Ihrig, Colin J.: "Remote AJAX Calls Using JSONP." cjihrig.com/blog/, 2012. Accessed: January 31, 2014.
http://cjihrig.com/blog/remote-ajax-calls-using-jsonp/

INSPIRE: "Directive 2007/2/EC of the European Parliament and of the Council." 2007.
http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2007:108:0001:0014:en:PDF

INSPIRE Consolidation Team: "INSPIRE Network Services Performance Guidelines." Technical report, European Commission, 2007.
http://inspire.jrc.ec.europa.eu/reports/ImplementingRules/network/Network_Services_Performance_Guidelines_%20v1.0.pdf

INSPIRE Network Services Drafting Team: "INSPIRE Technical Architecture – Overview (Version 1.2)." Technical report, European Commission, 2007.
http://inspire.jrc.ec.europa.eu/reports/ImplementingRules/network/INSPIRETechnicalArchitectureOverview_v1.2.pdf

INSPIRE Network Services Drafting Team: "INSPIRE Network Services Architecture (Version 3.0)." Technical report, European Commission, 2008.
http://inspire.jrc.ec.europa.eu/reports/ImplementingRules/network/D3_5_INSPIRE_NS_Architecture_v3-0.pdf

INSPIRE Network Services Drafting Team: "Draft Implementing Rules for Download Services (Version 3.0)." Technical report, European Commission, 2009.
http://inspire.jrc.ec.europa.eu/documents/Network_Services/INSPIREDrafImplementingRulesDownloadServices(version3.0).pdf

ISO: "Information Technology — Open Distributed Processing — Reference Model: Architecture." Technical report, International Organization for Standardization, 1996a.
http://standards.iso.org/ittf/PubliclyAvailableStandards/s020697_ISO_IEC_10746-3_1996(E).zip

ISO: "Information Technology — Open Distributed Processing — Reference Model: Foundations." Technical report, International Organization for Standardization, 1996b.
http://standards.iso.org/ittf/PubliclyAvailableStandards/s018836_ISO_IEC_10746-2_1996(E).zip

ISO: "Information Technology — Open Distributed Processing — Reference Model: Architectural Semantics." Technical report, International Organization for Standardization, 1998a.

http://standards.iso.org/ittf/PubliclyAvailableStandards/c020698_ISO_IEC_10746-4_1998(E).zip

ISO: "Information Technology — Open Distributed Processing — Reference Model: Overview." Technical report, International Organization for Standardization, 1998b. http://standards.iso.org/ittf/PubliclyAvailableStandards/c020696_ISO_IEC_10746-1_1998(E).zip

ISO: "ISO 19119: Geographic information – Services." 2001.

ISO: "ISO 19101:2002 Geographic information – reference model." 2002a.

ISO: "ISO 19115: Geographic information – Metadata." 2002b.

ISO: "Information Technology — Open Distributed Processing — Use of UML for ODP system specifications." Technical report, International Organization for Standardization, 2009. http://www.lcc.uma.es/%7Eav/download/UML4ODP_IS_V2.pdf

Jepsen, Thomas C.: "Just What Is an Ontology, Anyway?" In: *IT Professional*, volume 11 (5): pp. 22–27, 2009. ISSN 1520-9202.

Kaika, Maria: "The Water Framework Directive: A New Directive for a Changing Social, Political and Economic European Framework." In: *European Planning Studies*, volume 11 (3): pp. 299–316, 2003. http://www.tandfonline.com/doi/abs/10.1080/09654310303640

Klyne, Graham & Carroll, Jeremy J.: "Resource Description Framework (RDF): Concepts and Abstract Syntax." W3C recommendation, World Wide Web Consortium, 2004. http://www.w3.org/TR/2004/REC-rdf-concepts-20040210/

Kohlus, Jörn & Reimers, Hans-Christian: "Neue Herausforderungen im Datenmanagement für das europäische Meeresmonitoring - Das Projekt MDI-DE - Marine Daten-Infrastruktur in Deutschland." In: *Coastline Reports*, volume 16: pp. 115–126, 2010. http://databases.eucc-d.de/files/documents/00001004_Seiten%20aus%20CR16_115-126.pdf

Kontio, Mikko: "Architectural manifesto: Designing software architectures, Part 5: Introducing the 4+1 view model." IBM developerWorks, 2005. Accessed: January 31, 2014. http://de.scribd.com/doc/84656947/Architectural-manifesto-Designing-software-architectures-Part-5

Korduan, Peter: "Abbildung von marinen Geodaten durch INSPIRE Datenspezifikationen." In: "Geoinformationen für die Küstenzone Band 4," (edited by Karl-Peter Traub; Jörn Kohlus & Thomas Lüllwitz), pp. 33 – 43. 2013.

Korduan, Peter & Rüh, Christian: "Referenzmodell zum Aufbau einer marinen Geodateninfrastruktur in Deutschland (MDI-DE)." Technical report, Rostock University, 2013.

Kresse, Wolfgang; Danko, David M. & Fadaie, Kian: "Cultural and Linguistic Adaptability." In: "Springer Handbook of Geographic Information," (edited by Wolfgang Kresse & David M. Danko), chapter 13 Standardization. Springer Berlin Heidelberg, 2012. http://link.springer.com/book/10.1007/978-3-540-72680-7

Kruchten, Philippe: "Architectural Blueprints – the "4+1" View Model of Software Architecture." In: *IEEE Software*, volume 12 (6): pp. 42–50, 1995. ftp://public.dhe.ibm.com/software/rational/web/whitepapers/2003/Pbk4p1.pdf

Kuhn, Werner; Basedow, Sebastian; Brox, Christoph; Riedemann, Catharina; Rossol, Holger; Senkler, Kristian & Zens, Katharina: "Reference Model 3.0 – Geospatial Data Infrastructure (GDI) North-Rhine-Westfalia." Technical report, 2000. http://redgeomatica.rediris.es/metadatos/publica/Reference_GDI_NRW.pdf

Labonte, Jeff; Corey, Mark & Evangelatos, Tim: "Canadian Geospatial Data Infrastructure (CGDI) - Geospatial Information for the Knowledge Economy." In: *Geomatica*, volume 52: pp. 214–222, 1998.

cgdi.gc.ca/publications/General_information/geomatica_cgdi_E.pdf

Lakhani, Karim R.; Austin, Robert D. & Yi, Yumi: "Data.gov." Case study, Harvard Business School, 2010.
http://www.data.gov/sites/default/files/attachments/hbs_datagov_case_study.pdf

Lehfeldt, Rainer & Melles, Johannes: "Die Marine Dateninfrastruktur Deutschland MDI-DE." In: "Geoinformationen für die Küstenzone," (edited by Karl-Peter Traub; Jörn Kohlus & Thomas Lüllwitz), 3, pp. 107–116. Points Verlag, 2011.
http://projekt.mdi-de.org/images/mdi-de/Publikationen/Artikel_Beitraege/20111107110415418.pdf

Lekkerkerk, Huibert-Jan; Woolf, Andrew; Barrot, Dolors; Coley, Alexander; Eriksson, Helen; Fretter, Klaus; Hopfstock, Anja & Parslow, Peter: "INSPIRE Data Specification on Hydrography Version 3.0.1." 2010.
http://inspire.jrc.ec.europa.eu/documents/Data_Specifications/INSPIRE_DataSpecification_HY_v3.0.1.pdf

Lemon, David; Cox, Simon J. D.; Walker, Gavin; Atkinson, Robert A.; Fitch, Peter & O'Hagan, Rochelle G.: "The Water Resources Observation Network – Reference Model - Version 0.1." Technical report, CSIRO, 2007.
http://www.clw.csiro.au/publications/waterforahealthycountry/2007/wfhc-WRON-Reference-Model-V0.1.pdf

Lihteneger, Darja; Wilson, Debbie; Bergheim, Stein Runar; Borsa, Maciej; Coley, Alex; Gül, Ali; Haarman, Rob; Longhorn, Roger; Guilbert, Tom; Øverli, Tor Gunnar; Viarengo, Luca; Yüksel, Ebubekir & Lutz, Michael: "Data Specification on Area management/restriction/regulation zones and reporting units Version 3.0.0 rc3." 2013.
http://inspire.jrc.ec.europa.eu/documents/Data_Specifications/INSPIRE_DataSpecification_AM_v3.0rc3.pdf

Lord, Phillip: "Components of an Ontology." http://ontogenesis.knowledgeblog.org/514, 2010. Accessed: January 31, 2014.
http://ontogenesis.knowledgeblog.org/514

Lumb, Chris M; Fowler, Sarah L; Atkins, Stephen M; Gilliland, Paul M & Vincent, Malcolm A: "The Irish Sea Pilot: Developing marine nature conservation objectives for the Irish Sea." Technical report, Report to Defra by the Joint Nature Conservation Committee, Peterborough, 2004.
http://jncc.defra.gov.uk/PDF/objectives.pdf

Lupp, Markus: "Open Geospatial Consortium." In: "Encyclopedia of GIS," p. 815. Springer US, 2008. ISBN 978-0-387-30858-6.
http://link.springer.com/referenceworkentry/10.1007/978-0-387-35973-1_918

Maali, Fadi; Cyganiak, Richard & Peristeras, Vassilios: "Enabling Interoperability of Government Data Catalogues." In: "Electronic Government," (edited by Maria. Wimmer; Jean-Loup Chappelet; Marijn Janssen & Hans Scholl), volume 6228 of *Lecture Notes in Computer Science*, pp. 339–350. Springer Berlin Heidelberg, 2010. ISBN 978-3-642-14798-2.
http://dx.doi.org/10.1007/978-3-642-14799-9_29

Maali, Fadi; Cyganiak, Richard & Peristeras, Vassilios: "Re-using Cool URIs: Entity Reconciliation Against LOD Hubs." In: "Proceedings of the Linked Data on the Web Workshop 2011 (LDOW2011)," 2011.
http://events.linkeddata.org/ldow2011/papers/ldow2011-paper11-maali.pdf

Manola, Frank & Miller, Eric: "RDF Primer." W3c recommendation, W3C, 2004.
http://www.w3.org/TR/rdf-primer/

McGranaghan, Matthew: "Geospatial Data Infrastructure: Concepts, Cases, and Good Practice." In: *Annals of the Association of American Geographers*, volume 93 (1): pp. 257–259, 2003. ISSN 1467-8306.
http://dx.doi.org/10.1111/1467-8306.93126

McKeague, Peter; Corns, Anthony & Shaw, Robert: "Developing a Spatial Data Infrastructure for Archaeological and Built Heritage." In: *International Journal of Spatial Data Infrastructures Research*, volume 7: pp. 38–65, 2012.

Michaelis, Christopher D. & Ames, Daniel P.: "Web Feature Service (WFS) and Web Map Service (WMS)." In: "Encyclopedia of GIS," pp. 1259–1261. Springer US, 2008. ISBN 978-0-387-30858-6. http://dx.doi.org/10.1007/978-0-387-35973-1_459

Miles, Alistair: "SKOS: requirements for standardization." In: "DCMI '06 Proceedings of the 2006 international conference on Dublin Core and Metadata Applications: metadata for knowledge and learning," ACM, 2006. ISBN 970-692-268-7. http://dcpapers.dublincore.org/index.php/pubs/article/viewPDFInterstitial/839/835

Miles, Alistair & Bechhofer, Sean: "SKOS Simple Knowledge Organization System Reference." W3c recommendation, W3C, 2009. http://www.w3.org/TR/skos-reference

Miles, Alistair & Brickley, Dan: "SKOS Core Guide." W3C working draft, W3C, 2005. Published online on November 2nd, 2005 at http://www.w3.org/TR/2005/WD-swbp-skos-core-guide-20051102/. http://www.w3.org/TR/2005/WD-swbp-skos-core-guide-20051102/

Miles, Alistair & Brickley, Dan: "SKOS Simple Knowledge Organization System eXtension for Labels." W3c recommendation, W3C, 2009. http://www.w3.org/TR/skos-reference/skos-xl.html

Miles, Alistair; Matthews, Brian; Wilson, Michael & Brickley, Dan: "SKOS core: Simple Knowledge Organisation for the Web." In: "Proceedings of the 2005 international conference on Dublin Core and metadata applications," pp. 1–9. Dublin Core Metadata Initiative, 2005. http://dcpapers.dublincore.org/pubs/article/view/798

Millard, Keiran; Lowe, Dominic; Jensen, Hans Mose; Roesbeke, Marc; Brandini, Carlo; Hermida, Nuria; Tello, Maria Olvido; Sarretta, Alessandro; Reznik, Tomas; Lutz, Michael & Cetl, Vlado: "Data Specification on Oceanographic Geographical Features Version 3.0.0 rc3." 2013a. http://inspire.jrc.ec.europa.eu/documents/Data_Specifications/INSPIRE_DataSpecification_OF_v3.0rc3.pdf

Millard, Keiran; Lowe, Dominic; Jensen, Hans Mose; Roesbeke, Marc; Brandini, Carlo; Hermida, Nuria; Tello, Maria Olvido; Sarretta, Alessandro; Reznik, Tomas; Lutz, Michael & Cetl, Vlado: "Data Specification on Sea Regions Version 3.0.0 rc3." 2013b. http://inspire.jrc.ec.europa.eu/documents/Data_Specifications/INSPIRE_DataSpecification_SR_v3.0rc3.pdf

Mitchell, Rebecca; Williamson, Ann & O'Connor, Rod: "The development of an evaluation framework for injury surveillance systems." In: *BMC Public Health*, volume 9 (1): p. 260, 2009. ISSN 1471-2458. http://www.biomedcentral.com/1471-2458/9/260

Moltmann, Tim; Proctor, Roger; Hill, Katy & McGowen, Marian: "The Integrated Marine Observing System - delivering data-streams to support marine research and applications." In: "Proceedings of OCEANS 2010 IEEE, Sydney," IEEE, 2010.

MSFD: "Directive 2008/56/EC (Marine Strategy Framework Directive)." 2008. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2008:164:0019:0040:EN:PDF

Müller, Markus; Remke, Albert & Voges, Uwe: "14 – Katalogdienste und Metainformation." In: "Geodateninfrastruktur – Grundlagen und Anwendungen," (edited by Lars Bernard; Jens Fitzke & Roland M. Wagner), pp. 126–133. Wichmann, Heidelberg, 2004.

Müller, Uli & Mandery, Nico: "Leistungsfähigkeit aktueller WMS-Dienste." Presentation at Intergeo 2009, 2009. http://www.fossgis.de/w/images/9/9d/Geops_wmsperformance_090922.pdf

Müller, Uli & Mandery, Nico: "Qualitätsaspekte aktueller Web Map Services (WMS) als Grundlage einer dienstbasierten Geodateninfrastruktur." In: "Angewandte Geoinformatik 2010: Beiträge zum 22. AGIT-Symposium Salzburg," (edited by Josef Strobl; Thomas Blaschke &

Gerald Griesebner), volume XVIII, pp. 813 – 816. Wichmann, Berlin, 2010. https://www.vde-verlag.de/proceedings-en/537495104.html

Nairn, Alister D.: *Coastal and Marine Geospatial Technologies*, chapter 2 - The Development of an Australian Marine Spatial Information System (AMSIS) to Support Australian Government Ocean Policy and Multi-Use Marine Activities, pp. 17–29. Springer, 2009. www.ga.gov.au/image_cache/GA7900.pdf

Najar, Christine; Rajabifard, Abbas; Williamson, Ian P. & Giger, Christine: *Research and Theory in Advancing Spatial Data Infrastructure Concepts*, chapter A Framework for Comparing Spatial Data Infrastructures on the basis of Web Services and Metadata Management: An Australian-Swiss Case Study, pp. 201–213. 2007.

Najar, Christine Ruth: *A model-driven approach to management of integrated metadata-spatial data in the context of spatial data infrastructures*. Ph.D. thesis, ETH Zürich, 2006. http://e-collection.library.ethz.ch/eserv/eth:28733/eth-28733-02.pdf#search=%22(keywords_en:LIBRARIANSHIP)%22

Nasser, Hussein: "ArcGIS Server Performance." Blog post, 2009. Accessed: January 31, 2014. http://www.husseinnasser.com/2009/02/esri-arcgis-server-performance.html

Nebert, Douglas & Anthony, Michelle: "Spatial Data Infrastructure Concepts and Components." Presentation at Global Geospatial Conference (GSDI12), 2010. http://www.gsdi.org/gsdiconf/gsdi12/slides/ws1.pdf

Nebert, Douglas D.: *Developing Spatial Data Infrastructures: The SDI Cookbook*. GSDI, 2004. http://www.gsdi.org/docs2004/Cookbook/cookbookV2.0.pdf

Nebert, Douglas D.: "Introduction to Geospatial Web Services." Workshop presentation at Global Geospatial Conference (GSDI11), 2009. http://www.gsdi.org/gsdi11/wrkshpslides/w1.1c.pdf

NRC: "Marine geospatial data - safeguarding Canada's coastal and offshore areas." Monograph, Natural Resources Canada, 2003. www.cgdi.gc.ca/programsCommittees/proCom_marine/marine_factsheet_E.pdf

Oceans Portal Project Governance Working Group: "Oceans Portal Project - Governance Working Group Discussion Paper." Technical report, AAD, NOO, CMAR, BoM, AIMS, GA, 2006. http://www.aodc.gov.au/files/GovernanceDiscussionPaper.pdf

O'Dea, Elizabeth; Cummins, Valerie & Dwyer, Edward: "Developing an Informational Web Portal for Coastal Data in Ireland: Data Issues in the Marine Irish Digital Atlas." In: "Proceedings of Oceanology International," 2004a. http://cmrc.ucc.ie/publications/conf_presentations/oceanology2004_mida.pdf

O'Dea, Elizabeth; Cummins, Valerie; Wright, Dawn; Dwyer, Edward & Ameztoy, Iban: "Report on Coastal Mapping and Informatics Trans-Atlantic Workshop 1: Potentials and Limitations of Coastal Web Atlases." Technical report, University College Cork, Ireland, Coastal & Marine Resources Centre, 2007.

O'Dea, Elizabeth; Dwyer, Edward; Cummins, Valerie; Dunne, Declan & Tuama, Eamonn Ó: "Improving Access to Coastal Information: Metadata in the Marine Irish Digital Atlas." In: "Proceedings of ECO-IMAGINE," 2004b. http://www.gisig.it/eco-imagine/full%20papers/O'Dea.PDF

O'Dea, Elizabeth; Dwyer, Edward; Cummins, Valerie; Giménez, Dídac Perales Í & Dunne, Declan: *Coastal and Marine Geospatial Technologies*, chapter 4 - Harmonising Marine Information Exchange in Ireland, pp. 41 – 50. Springer, 2009.

OGC: "TEAM Engine – Getting Started Guide." Technical report, The Open Geospatial Consortium, 2013. http://download2.polytechnic.edu.na/pub4/sourceforge/t/te/teamengine/Team%20Engine/4.0beta1/teamengine-4.0-beta1.pdf

O'Hagan, Rochelle G.; Atkinson, Robert A.; Cox, Simon J. D.; Fitch, Peter; Lemon, David & Walker, Gavin: "A Reference Model for a Water Resources Observation Network." In: "MODSIM 2007 –

International Congress on Modelling and Simulation," (edited by Les Oxley & Don Kulasiri), pp. 1145–1151. Modelling and Simulation Society of Australia and New Zealand, 2007.
http://mssanz.org.au/MODSIM07/papers/18_s56/AReferenceModel_s56_O%27Hagan_.pdf

OMG: "OMG Unified Modeling Language (OMG UML), Superstructure, Version 2.4.1." 2011.
http://www.omg.org/spec/UML/2.4.1

Parasuraman, A.; Zeithaml, Valarie A. & Berry, Leonard L.: "A Conceptual Model of Service Quality and Its Implications for Future Research." In: *Journal of Marketing*, volume 49 (4): pp. pp. 41–50, 1985. ISSN 00222429.
http://www.jstor.org/stable/1251430

Percivall, George: "OGC Reference Model Version 0.1.3." Technical report, Open Geospatial Consortium, 2003.
https://portal.opengeospatial.org/files/?artifact_id=3836

Peter, Wolfgang: "Korrelationskoeffizient." Blog, 2006. Accessed: January 31, 2014.
http://arbeitszimmer.statistik-peter.at/2006/08/24/korrelationskoeffizient/

Proctor, Roger; Roberts, Kate & Ward, Brendon J.: "A data delivery system for IMOS, the Australian Integrated Marine Observing System." In: *Advances in Geosciences*, volume 28: pp. 11–16, 2010.
http://www.adv-geosci.net/28/11/2010/

Raimundo, F. Dos Santos Jr. & Chang-Tien, Lu: "Geography Markup Language (GML)." In: "Encyclopedia of GIS," pp. 364–368. Springer US, 2008. ISBN 978-0-387-30858-6.

Rajabifard., Abbas & Williamson, Ian P.: "Spatial data infrastructures: concept, SDI hierarchy and future directions." 2001.
http://eprints.unimelb.edu.au/archive/00001151/

Reed, Greg; Michailov, Nick & Pissierssens, Peter: "The IODE Ocean Data Portal: Concept Paper." Technical report, INTERGOVERNMENTAL OCEANOGRAPHIC COMMISSION (of UNESCO), 2007.
http://www.oceandocs.net/bitstream/1834/2770/1/IODE-XIX_15.pdf

Rice, Jake C.; Trujillo, Valentin; Jennings, Simon; Hylland, Ketil; Hagstrom, Olle; Astudillo, Armando & Jensen, Jørgen Nørrevang: "Guidance on the Application of the Ecosystem Approach to Management of Human Activities in the European Marine Environment." Technical report, ICES, Copenhagen, Denmark, 2005.
http://www.ices.dk/products/crrs/EMS%20Guidancefinal.pdf

Robinson, Catherine J.; Taylor, Bruce M.; eonie Pearson a; O'Donohue, Mark & Harman, Ben: "A SMART assessment of water quality partnership needs in Great Barrier Reef catchments." In: *Australasian Journal of Environmental Management*, volume 16 (2): pp. 84–93, 2009.
http://www.tandfonline.com/doi/abs/10.1080/14486563.2009.9725222

Roscher, Sabine; Schmidt, Anne; Ssymank, Axel; Tirry, Diederik; Valland, Nils; Vischer-Leopold, Mareike; Wilkinson, Steve; Tuchyna, Martin & Abramic, Andrej: "Data Specification on Bio-geographical Regions Version 3.0.0 rc3." 2013.
http://inspire.jrc.ec.europa.eu/documents/Data_Specifications/INSPIRE_DataSpecification_BR_v3.0rc3.pdf

Rossmanith, Thomas & Schupp, Anja: "8. Rahmenpapier GDI-Südhessen – Organisations-, Rollen- und Finanzierungsmodell." Technical report, Hessische Verwaltung für Bodenmanagement und Geoinformation, 2008.
http://www.gdi-suedhessen.de/fileadmin/user_upload/ikgis/GDI_Sued/Aktuelles_Informationen/Rahmenpapiere/GDI_Suedhessen_RP_08_Organisations_und_Rollenmodell.pdf

Rueda, Carlos; Bermudez, Luis & Fredericks, Janet: "The MMI Ontology Registry and Repository: A portal for Marine Metadata Interoperability." In: "OCEANS 2009, MTS/IEEE Biloxi - Marine Technology for Our Future: Global and Local Challenges," pp. 1 –6. 2009.
http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5422206&tag=1

Russell, Ian: "Hydrography and marine spatial data infrastructure - Programme

and Presentation Abstracts." Brochure, The Hydrographic Society UK, 2009. www.ths.org.uk/documents/ths.org.uk/downloads/symposium_brochure.pdf

Salgé, François; Quak, Wilko; Mikkola, Aaro; van Noordt, Anneloes; Hazeu, Gerard Willem; Goñi, Isabel; Hernández, Julián Delgado; Krause, Kai-Uwe; Andrzejewska, Maria; Kasanko, Marjo; Richter, Walter & deLima, Maria Vanda Nunes: "Data Specification on Land Use Version 3.0.0 rc3." 2013. http://inspire.jrc.ec.europa.eu/documents/Data_Specifications/INSPIRE_DataSpecification_LU_v3.0rc3.pdf

Schandl, Thomas: "Transforming spreadsheets into SKOS with Google Refine." blog.semantic-web.at, 2011. Accessed: January 31, 2014. http://blog.semantic-web.at/2011/02/17/transforming-spreadsheets-into-skos-with-google-refine/

Schandl, Thomas & Blumauer, Andreas: "PoolParty: SKOS Thesaurus Management Utilizing Linked Data." In: "The Semantic Web: Research and Applications," (edited by Lora Aroyo; Grigoris Antoniou; Eero Hyvönen; Annette Teije; Heiner Stuckenschmidt; Liliana Cabral & Tania Tudorache), volume 6089 of *Lecture Notes in Computer Science*, pp. 421–425. Springer Berlin Heidelberg, 2010. ISBN 978-3-642-13488-3. http://dx.doi.org/10.1007/978-3-642-13489-0_36

Schut, Peter: "CGDI FRAMEWORK DATA DEFINITION - A working definition prepared by the GeoConnections Framework Data Node." 2001. Accessed: January 31, 2014. http://www.geos.ed.ac.uk/~gisteac/proceedingsonline/Source%20Book%202004/SDI/National/Canada/Canadaian%20GDI%20Framework%20Data%20Definition.pdf

Seeley, Becky: "The elements of the MEDIN discovery standard." In: "Marine Data Standards Workshop 25 & 26th October 2011 (Liverpool and Isle of Man)," 2011. http://www.oceannet.org/marine_data_standards/medin_approved_standards/documents/medinworkshop.ppt

Serrano, Jean-Jacques; Laxton, John; Ash, Kristine; Batalla, Xavier Berástegui; Bergman, Stefan; Cassard, Daniel; Follestad, Bjørn; Hughes, Andrew; Larsen, Uffe; Nalecz, Tomasz; Pen, Simon; Sorés, László; Vuollo, Jouni & Tomas, Robert: "Data Specification on Mineral Resources Version 3.0.0 rc3." 2013. http://inspire.jrc.ec.europa.eu/documents/Data_Specifications/INSPIRE_DataSpecification_MR_v3.0rc3.pdf

Sherin, Andrew G.: "COINAtlantic: Planning for a Sustainable Future." In: "Proceedings of CoastGIS 2007: 8th International Symposium on GIS and Computer Mapping for Coastal Zone Management," 2007. www.gioc.unican.es/CoastGIS07/files/Sherin_CoastGIS07_paper.doc

Sherin, Andrew G.; Butler, Michael J.A.; Leblanc, Claudette; Gillespie, Randy & Collins, Norval: *Coastal and Marine Geospatial Technologies*, chapter 8 - Coastal Ocean Information Network (ATLANTIC): From Concept to Reality: A Status Report, pp. 73–85. Springer, 2009. http://www.springerlink.com/content/rx606r3q1g5p6462/

Sinha, Abhinaya: "Web Feature Service (WFS)." In: "Encyclopedia of GIS," pp. 1256–1259. Springer US, 2008. ISBN 978-0-387-30858-6. http://dx.doi.org/10.1007/978-0-387-35973-1_459

Sotis, Gianni: "Testing Services – Engineering ReportGEOSS Architecture Implementation Pilot, Phase 2." Report, GEO Architecture Implementation Pilot, 2009. Part 2. http://www.ogcnetwork.net/system/files/AIPTestingServices_ER_Final_1_0_pages_38to74.pdf

Staub, Peter: *Über das Potenzial und die Grenzen der semantischen Interoperabilität von Geodaten*. Ph.D. thesis, ETH Zürich, 2009. http://e-collection.library.ethz.ch/eserv/eth:41507/eth-41507-02.pdf

Staveley, Alex: "The "4+1" View Model of Software Architecture." 2011. Accessed: January 31, 2014. http://java.dzone.com/articles/%E2%80%9C41%E2%80%9D-view-model-software

Steudler, Daniel; Rajabifard, Abbas & Williamson, Ian P.: *A Multi-View Framework to Assess Spatial Data Infrastructures*, chapter 10 - Evaluation and Performance Indicators to Assess Spatial Data Infrastructure Initiatives, pp. 193–210. 2008. http://www.csdila.unimelb.edu.au/publication/books/mvfasdi/MVF_assessment_SDI.pdf

Strain, Lisa: *An SDI Model to Include the Marine Environment*. Master's thesis, Department of Geomatics Faculty of Engineering University of Melbourne, 2006. http://www.csdila.unimelb.edu.au/publication/thesis/Lisa_Strain_Msc_Thesis.pdf

Strain, Lisa; Rajabifard, Abbas & Williamson, Ian P.: "Marine Administration and Spatial Data Infrastructures." In: *Marine Policy*, 2006. http://www.sciencedirect.com/science/article/pii/S0308597X05000345

Swanton, Phillipa: "MAGIC INFORMATION NOTE - Data Standard." Technical report, Department for Environment, Food and Rural Affairs (DEFRA), 2001. http://magic.defra.gov.uk/ProjectLibrary/InfoNotes/magi0015.pdf

Tóth, Katalin; Portele, Clemens; Illert, Andreas; Lutz, Michael & de Lima, Maria Nunes: "A Conceptual Model for Developing Interoperability Specifications in Spatial Data Infrastructures." Technical report, European Commission, Joint Research Centre, Institute for Environment and Sustainability, 2012. http://inspire.jrc.ec.europa.eu/documents/Data_Specifications/IES_Spatial_Data_Infrastructures_(online).pdf

Tuchyna, Martin; Tirry, Diederik; Casado, Pablo Burgos; Chudy, Radoslav; Kristensen, Søren Elkjær; Remmelts, Gijs; Rhodes, Chris; Semrad, Peter; Szabo, Marta; Suri, Marcel & Thomas, Lisa: "Data Specification on Energy Resources Version 3.0.0 rc3." 2013. http://inspire.jrc.ec.europa.eu/documents/Data_Specifications/INSPIRE_DataSpecification_ER_v3.0rc3.pdf

Vachon, Don; Lafond, Pierre & Evangelatos, Timothy V.: "Delivering Marine Geospatial Data on the Web - The Canadian Department of Fisheries and Oceans' GeoPortal." In: *International Hydrographic Review*, volume 3 (3): pp. 63–71, 2002. http://journals.hil.unb.ca/index.php/ihr/article/download/18620/20299

Vaez, Sheelan & Rajabifard, Abbas: "Seamless SDI Design by Using UML Modelling." In: *International Journal of Spatial Data Infrastructures Research*, volume 7: pp. 207–224, 2012. http://ijsdir.jrc.ec.europa.eu/index.php/ijsdir/article/view/282/328

Vallecillo, Antonio: *DINTEL Edition on Software Engineering*, chapter RM-ODP: The ISO Reference Model for Open Distributed Processing, pp. 69–99. 3. 2001. http://www.lcc.uma.es/~av/Publicaciones/00/odpeng.pdf

Vessie, David; van de Poll, Robert; Nichols, Susan & Monahan, David: "GIS mapping and analysis tools for use within the coastal zone." In: "Proceedings of the 2nd Biennial Coastal GeoTools Conference, Charleston, SC, January 8-11," 2001.

Vretanos, Panagiotis A.: "OpenGIS Web Feature Service 2.0 Interface Standard." 2010.

Welle-Donker, Frederika: "Successful SDIs: does the Marine Geo-sector provide a much-needed beacon?" In: "Proceedings of the GSDI 12 World Conference," 2010. http://repository.tudelft.nl/view/ir/uuid%3A5daf1957-4464-44f6-8182-daacd832102c/

Williamson, Ian P.; Rajabifard, Abbas & Binns, Andrew: "Issues in developing marine SDI." In: "Proceedings of International Workshop on Administering the Marine Environment – The Spatial Dimensions," 2004. http://www.csdila.unimelb.edu.au/publication/conferences/Issues%20in%20developing%20Marine%20SDI.pdf

Wood, Louisa: "Global Marine Protection Targets: How S.M.A.R.T are They?" In: *Environmental Management*, volume 47 (4): pp. 525–535, 2011.

http://www.springerlink.com/content/7h3525827k40w633/?MUD=MP

WS6 Communications: "Managing UK marine data." In: *The Marine Scientist*, volume 32 (32): pp. 14–16, 2010. http://www.oceannet.org/library/publications/documents/marine_scientist_article.pdf

Wytzisk, Andreas: "Organizational Setup of the Saxon SDI." In: "3rd INSPIRE Training," con terra, 2012. http://www.inspiration-westernbalkans.eu/5/9/5/3/7/7/Organizational_Setup_of_the_Saxon_SDI.pdf

Yang, Chaowei Phil & Evans, John: "Network GIS Performance." In: "Encyclopedia of GIS," (edited by Shashi Shekhar & Hui Xiong), pp. 787–790. Springer, 2008. ISBN 978-0-387-30858-6. http://dblp.uni-trier.de/db/reference/gis/gis2008.html#YangE08

Zampoukas, Nikolaos; Piha, Henna; Bigagli, Emanuele; Hoepffner, Nicolas; Hanke, Georg & Cardoso, Ana Cristina: "Monitoring for the Marine Strategy Framework Directive: Requirements and Options." Report, European Commission Joint Research Centre, 2012. http://publications.jrc.ec.europa.eu/repository/bitstream/111111111/23169/1/lbna25187enn.pdf

Acknowledgements

This thesis would not have been possible without all the people who helped and supported me throughout it. Special thanks go to my supervisor Prof. Dr.-Ing. Ralf Bill. His extensive experience and input were invaluable in the realization of this thesis. I am also very thankful to my examiners Dr.-Ing. Rainer Lehfeldt from the German Federal Waterways Engineering and Research Institute and Prof. Dr. Christophe Claramunt from the French Naval Academy Research Institute for their support, continued assistance and guidance.

I would also like to thank the past and present members of the MDI-DE project, the sub project directors Johannes Melles (German Federal Maritime and Hydrographic Agency, BSH) and Peter Hübner (German Federal Agency for Nature Conservation, BfN) are mentioned representatively at this point.

Furthermore I would like to acknowledge the generous assistance of my colleagues at the Professorship for Geodesy and Geoinformatics at the University of Rostock. In particular I would like to thank Dr.-Ing. Peter Korduan. His depth of knowledge and devotion to technical research created a positive working atmosphere for me. Dr.-Ing. Görres Grenzdörffer was also helpful with his useful hints and advices.

Many thanks also go to the unknown reviewers of my papers for their advice and hints.

My appreciation and thanks also goes to my fiancée Sarah Seip for her continuous support and love. Without her this thesis would have not been possible. In conjunction with this I would also like to thank Prof. Dr. med. H. W. S. Schroeder and his colleagues from the Greifswald University Hospital.