

Entwurf eines allgemeinen Modells der raumbezogenen Metainformationsverarbeitung

Dissertation

zur

Erlangung des Doktorgrades (Dr. rer. nat.)

der

Mathematisch-Naturwissenschaftlichen Fakultät

der

Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

Dipl.-Geogr. Markus Müller

aus

Karlsruhe

Bonn, April 2002

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen
Fakultät der Rheinischen Friedrich-Wilhelm-Universität Bonn

Inhaltsverzeichnis

<u>INHALTSVERZEICHNIS</u>	<u>III</u>
--	-------------------

<u>ABBILDUNGSVERZEICHNIS</u>	<u>VII</u>
---	-------------------

<u>TABELLENVERZEICHNIS</u>	<u>IX</u>
---	------------------

<u>1</u>	<u>EINFÜHRUNG</u>	<u>1</u>
1.1	MOTIVATION	1
1.2	PROJEKTKONTEXT	2
1.3	STRUKTUR DER ARBEIT.....	3
<u>2</u>	<u>WISSENSCHAFTLICHE EINORDNUNG</u>	<u>6</u>
2.1	GEOGRAPHICAL INFORMATION SCIENCE.....	6
2.2	UMWELTINFORMATIK	8
2.3	SEMIOTIK	9
2.4	ZUSAMMENFASSUNG.....	11
<u>3</u>	<u>REFERENZRAHMEN UND METHODIK</u>	<u>12</u>
3.1	AUFBAU UND NATUR VON WISSENSCHAFT	12
3.2	ABGRENZUNG DES PROBLEMBEREICHES.....	14
3.3	THEORIE- UND MODELLBEGRIFF	17
3.4	BEZIEHUNG ZUM OPENGIS-KONTEXT	18
3.4.1	DAS INTEROPERABILITÄTSPROBLEM	18
3.4.2	RELEVANTE OGC-SPEZIFIKATIONEN.....	22
3.4.3	METAINFORMATION UND <i>OPENGIS</i>	24
3.4.4	ISO 19115 GEOGRAPHIC INFORMATION - METADATA.....	27

3.4.5	DIENTEDEFINITION BEI OGC UND ISO	30
3.4.6	OPENGIS KATALOGDIENSTE	30
3.5	ZUSAMMENHANG DES <i>OPENGIS</i> MODELLS UND DER SEMIOTISCHEN	
	THEORIE	35
3.6	VORGEHENSWEISE.....	36
<u>4</u>	<u>AKTUELLE ANSÄTZE ZUR</u>	
	<u>METAINFORMATIONSVERRARBEITUNG</u>	<u>37</u>
4.1	STAND DER FORSCHUNG	37
4.2	ANSÄTZE ZUR METAINFORMATIONSVERRARBEITUNG IN VERSCHIEDENEN	
	DISZIPLINEN	40
4.3	DATENBANKEN.....	41
4.4	DATA WAREHOUSE.....	42
4.5	UMWELTINFORMATIONSSYSTEME	43
4.5.1	UMWELTDATENKATALOG	44
4.5.2	UMWELTINFORMATIONSNETZ DEUTSCHLAND	47
4.5.3	EEA, EIONET & CDS.....	48
4.6	GEOINFORMATIONEN	49
4.6.1	CONTENT STANDARD FOR DIGITAL GEOSPATIAL METADATA	50
4.6.2	SPATIAL DATA TRANSFER STANDARD.....	51
4.6.3	ENV GEOGRAPHIC INFORMATION – DATA DESCRIPTION – METADATA	52
4.6.4	INGEO-MIS UND INGEO-MDF.....	53
4.6.5	GIS-POOL	54
4.6.6	GEODATENINFRASTRUKTUR NORDRHEIN-WESTFALEN	55
4.6.7	GEOMIS.BUND.....	55
4.7	DIGITALE BIBLIOTHEKEN	56
4.8	INTERNET	57
4.9	MEDIZININFORMATIK	60
4.10	ZUSAMMENFASSUNG.....	61
<u>5</u>	<u>EIN MODELL DER RAUMBEZOGENEN</u>	
	<u>METAINFORMATIONSVERRARBEITUNG</u>	<u>65</u>
5.1	BESCHREIBUNG DES ANSATZES	65
5.2	OBJEKTORIENTIERTE DENKWEISE	65

5.3	INFORMATIONEN UND DATEN	66
5.4	ZEICHEN UND KODE	68
5.5	DIE SPEZIELLE NATUR VON GEO-INFORMATIONEN	70
5.6	EXKURS: ARTEN VON RAUMBEZÜGEN.....	74
5.7	UMWELTINFORMATIONEN UND GEOINFORMATIONEN.....	75
5.8	METAINFORMATION UND METADATEN	76
5.9	RAUMBEZOGENE METAINFORMATION	87
5.10	ZUSAMMENFASSUNG.....	90
<u>6</u>	<u>DISKUSSION UND AUSBLICK.....</u>	<u>91</u>
6.1	NUTZEN EINER THEORIE DER RAUMBEZOGENEN METAINFORMATIONSVERRARBEITUNG.....	91
6.2	GEGENWÄRTIGE ENTWICKLUNGEN IM BEREICH DER RAUMBEZOGENEN METAINFORMATIONSVERRARBEITUNG.....	92
6.2.1	BEDEUTUNG VON XML.....	92
6.2.2	STANDARDISIERUNG VON METADATEN.....	93
6.2.3	INTEROPERABLE GEODATENKATALOGE DURCH UNTERSTÜTZUNG VON OGC –SPEZIFIKATIONEN	94
6.3	KONZEPTION UND ENTWICKLUNG EINES GEODATENKATALOGS AM BEISPIEL DES METAINFORMATIONSSYSTEMS VON LUXEMBURG	95
6.3.1	ALLGEMEINES	95
6.3.2	FUNKTIONSUMFANG.....	96
6.3.3	DATENMODELL	98
6.3.4	SYSTEMARCHITEKTUR	99
6.4	DIE ZUKUNFT DER RAUMBEZOGENEN METAINFORMATIONSVERRARBEITUNG.....	102
<u>7</u>	<u>ZUSAMMENFASSUNG</u>	<u>104</u>
<u>8</u>	<u>ANHANG: NOTATION VERWENDETER UML-DIAGRAMME</u>	<u>107</u>
8.1	KLASSE	107
8.2	ASSOZIATION	107
8.3	ROLLE.....	108

8.4	GENERALISIERUNG/SPEZIALISIERUNG	108
8.5	AGGREGATION.....	109
8.6	PAKETE.....	110
9	<u>LITERATUR.....</u>	<u>111</u>

Abbildungsverzeichnis

Abbildung 1. Struktur einer Wissenschaft (Quelle: nach Abler, Adams & Gould., 1971)	13
Abbildung 2. Ordnung von Erfahrungen (Quelle: nach Abler, Adams & Gould., 1971, verändert)	15
Abbildung 3. Neun Ebenen der Abstraktion (Quelle: OGC 1999a)	24
Abbildung 4. Beziehung zwischen Feature und Feature Collection (Quelle: nach OGC 1999a)	25
Abbildung 5. ISO 19115 Paketstruktur (Quelle: ISO 2000)	28
Abbildung 6. ISO 19115 Profile (Quelle: ISO 2000)	29
Abbildung 7. Architektur des Referenzmodells (Quelle: OGC 1999e)	32
Abbildung 8. Zerlegung der Katalogdienste (Quelle: OGC 1999e)	33
Abbildung 9. Katalog Objektmodell (Quelle: OGC 1999e)	34
Abbildung 10. Information	66
Abbildung 11. Diagramm der Klasse „Zoologische Art“	68
Abbildung 12. Zeichen	68
Abbildung 13. Ausschnitt aus der Benutzeroberfläche von GROMS	69
Abbildung 14. Geoinformation	71
Abbildung 15. Verteilung des Ozonmaximums in der BRD 1995 (Quelle: Senatsverwaltung für Stadtentwicklung, Umweltschutz und Technologie 1996)	72
Abbildung 16. Umweltinformation	76
Abbildung 17. Kommunikationsweg der Geoinformation	79
Abbildung 18. Rollen von Metainformation	85
Abbildung 19. Katalogeintrag als Zeichen	87
Abbildung 20. Geo-Metainformation	88
Abbildung 21. Metadatenmodell von MISLux in UML Notation	98
Abbildung 22. MISLux Architekturdiagramm	99
Abbildung 23. Klasse	107
Abbildung 24. Assoziation	107
Abbildung 25. Rollen	108

Abbildung 26. Generalisierung/Spezialisierung	108
Abbildung 27. Aggregation	109
Abbildung 28. Paket	110

Tabellenverzeichnis

Tabelle 1. Überblick über genannte Metadaten-Initiativen.....	62
Tabelle 2. Zoologische Arten	67

1 Einführung

1.1 Motivation

Im Allgemeinen werden Daten nicht von den selben Personen erfasst und erzeugt, welche sie nutzen. Häufig ist bei der Erstellung der Daten der zukünftige Benutzungskontext nicht bekannt oder zum Zeitpunkt der Erstellung noch nicht einmal existent. Hieraus resultiert der Bedarf an Informationen, die es erlauben über die Relevanz von Datenbeständen für bestimmte Fragestellungen zu entscheiden und welche die sinnvolle Nutzung von Daten ermöglichen. Eben dies ist die Aufgabe von Metainformationen.

Metainformationen dienen demnach der Identifizierung von Informationsressourcen und der Benutzung dieser Ressourcen außerhalb ihres Entstehungskontextes. Sie sind somit Mittel zur Werterhaltung von und Wertschöpfung mit Informationen. Da Geo- und Umweltdaten sehr heterogen und komplex sind und die Zusammenführung von Geodaten aus verschiedenen Quellen viele Probleme aufwirft, ist die Metainformationsverarbeitung eines der zentralen Themen in den Disziplinen *Geographical Information Science* sowie Umweltinformatik und dementsprechend im Blickfeld von Praxis und Wissenschaft. Umso bedauerlicher ist das Theoriedefizit, das in diesem Bereich bislang vorherrscht. Die gängige und immer wieder zitierte Definition für Metainformation/Metadaten lautet "Daten über Daten" (beispielsweise in GREVE 1996, STROBL 1995), welche pragmatisch ausreichend ist, aber genug Spielraum für gänzlich unterschiedliche Herangehensweisen lässt. Diese Uneinheitlichkeit hat ihre Begründung auch darin, dass Metainformation einen in verschiedenen Disziplinen behandelten Forschungsgegenstand darstellt, beispielsweise im digitalen Bibliothekswesen, der Umweltinformatik dem *Data Warehousing* und vor allem im Zusammenhang mit dem Internet.

Eine Folge dieser Uneinheitlichkeit ist die inhomogene Terminologie im Umgang mit Metainformation und die daraus folgenden Kommunikationsprobleme beim Neuaufbau von solchen Systemen zwischen Planern, Entwicklern und Nutzern.

Ziel der vorliegenden Arbeit ist es, eine fundierte und einheitliche Terminologie zu entwickeln, die die Grundlage für den Aufbau von raumbezogenen Metainformationssystemen darstellt. Das Kernproblem, das in dieser Arbeit gelöst werden soll, besteht darin mittels dieser einheitlichen Terminologie zu einem Vorschlag eines allgemeinen Konsenses über die Definition von Begriffen der Metainformationsverarbeitung zu gelangen. Dies erscheint notwendig in Hinsicht auf die vielfältigen Aufgaben, die Metainformationen im Bereich der raumbezogenen Informationsverarbeitung zu lösen imstande sind. Im Mittelpunkt der Anwendungsfälle, die den Hintergrund dieser Arbeit bilden, steht die Aufgabe der Identifizierung und des Zugriffs auf Geodaten durch Personen, welche diese Datensätze nicht selbst erzeugt haben. Das Werkzeug zur Lösung dieser Aufgaben stellen Geodatenkataloge dar. Hierfür ist es notwendig, eine Theoriebildung und Begriffsdefinition für raumbezogene Metainformationen durchzuführen, die den Ausgangspunkt für diese Terminologie darstellt. Ergebnis einer solchen Theoriebildung ist ein allgemeingültiges Denkgebäude der Geo-Metainformationsverarbeitung. Hierfür werden hauptsächlich Erkenntnisse aus der Semiotik und den theoretischen Grundlagen der Geo- und Umweltinformatik benutzt. Besonderes Gewicht wird auf eine mit der Denkweise des *OpenGIS Consortiums* (OGC) kompatible Modellbildung gelegt. Der Grund hierfür liegt in der großen Bedeutung, die das OGC als zentrale Norminstanz für den Bereich der praktischen raumbezogenen Informationsverarbeitung spielt.

Anhand eines praktischen Beispiels, dem Metainformationssystem für Geodaten des Großherzogtums Luxemburg (MISLux), werden die theoretischen und konzeptionellen Erkenntnisse erläutert. MISLux stellt ein auf den aktuellen Entwicklungen und Standards aufbauendes Katalogsystem dar, welches die in dieser Arbeit entwickelten Ergebnisse veranschaulicht und deren Umsetzung zeigt.

1.2 Projektkontext

Der Projektkontext für die vorliegende Arbeit ergibt sich aus der Beteiligung an den Projekten GIS-Pool (Ministerium für Raumordnung, Landwirtschaft

und Umwelt Sachsen-Anhalt), LANIS-Bund (Bundesamt für Naturschutz), FIS-Broker (Berliner Senatsverwaltung für Stadtentwicklung), dem Metainformationssystem für Geodaten des Bundes (GeoMIS.Bund, Bundesamt für Kartographie und Geodäsie), der Geodateninfrastruktur Nordrhein-Westfalen (GDI-NRW, Staatskanzlei Nordrhein-Westfalen) und nicht zuletzt dem Metainformationssystem für Geodaten des Großherzogtums Luxemburg im Rahmen des Projektverbundes EXSE (GIS EXperimental SErver) am Geographischen Institut der Universität Bonn. Alle diese Projekte haben gemeinsam, dass sie im Kern Metainformationssysteme sind oder diese beinhalten und dass die Geodatenverarbeitung eines ihrer Kernprobleme darstellt. Die in diesen Projekten gemachten Erfahrungen bilden den empirischen Hintergrund dieser Studie. Umgekehrt besteht der Anspruch, auch wenn sich diese Arbeit schwerpunktmäßig mit theoretischen Aspekten der *Geographical Information Science* befasst, die gewonnenen Erkenntnisse in die praktische Geo-Metainformationsverarbeitung zurückfließen zu lassen und den Aufbau von Geo-Metainformationssystemen zu erleichtern.

1.3 Struktur der Arbeit

Am Beginn der Arbeit befindet sich eine allgemeine und kurze Einführung in die Motivation und Struktur der Arbeit (das vorliegende Kapitel 1). Danach wird in Kapitel 2 der wissenschaftliche Hintergrund abgesteckt, der sich schwerpunktmäßig im Bereich der *Geographical Information Science* befindet, darüber hinaus aber wichtige Aspekte der Umweltinformatik enthält und vor allem die Brücke zur Semiotik schlägt. Diese Verbindung stellt die eigentliche wissenschaftliche Herausforderung der Arbeit dar, da zum Erreichen des gesetzten Ziels, die Entwicklung eines allgemeinen Modells der raumbezogenen Metainformationsverarbeitung, auf die Mittel der Semiotik zurückgegriffen wird.

Nachdem die wissenschaftliche Ansiedlung und der aktuelle Forschungsstand geklärt sind, wird in Kapitel 3 der Referenzrahmen beschrieben und die methodische Vorgehensweise erläutert. Hierbei werden einleitende Überlegungen zur Natur von Wissenschaft und wissenschaftlichem Arbeiten angestellt. Mittels eines einfachen epistemologischen Modells wird der

Kernbereich der Arbeit definiert und die Verbindung zwischen raumbezogener Informationsverarbeitung und Semiotik dargestellt. Die Definition des Referenzrahmens wird abgeschlossen durch die Vorstellung der abstrakten Konzepte des OGC und ihrer Bedeutung für diese Studie. Die Anwendung der *OpenGIS*-Konzepte wird deshalb vorgenommen, weil sie bereits die gesamte Geoinformationsverarbeitung stark beeinflussen und voraussichtlich in den nächsten Jahren noch stärker beeinflussen werden.

Die Darstellung des derzeitigen Standes von Forschung und Praxis ist die Aufgabe des darauf folgenden Kapitels 4. Dessen erster Teil (Abschnitt 4.1) befasst sich mit dem Stand der Forschung. Da diese Untersuchung in verschiedenen Teilen der Arbeit durchgeführt wird, besteht der Inhalt des Abschnittes 4.1 hauptsächlich aus Verweisen auf diese anderen Textstellen. Des weiteren wird in Kapitel 4 die Beziehung der Arbeit zu den Forschungsgebieten Ontologien und Wissensmanagement diskutiert. Unter Berücksichtigung des durch die Denkweise des OGC definierten Referenzrahmens werden in Abschnitt 4.2 bis 4.10 beispielhaft und ohne Anspruch auf Vollständigkeit Ansätze zur Metainformationsverarbeitung aus Sicht verschiedener Disziplinen vorgestellt und diskutiert. Dieses Kapitel dient vor allem dazu, dem Leser einen Einblick in die inhomogene Umgangsweise mit Metainformation und einen Hinweis auf spezielle raumbezogene Probleme in diesem Zusammenhang zu geben. Darüber hinaus werden hier die für die Modellbildung notwendigen Setzungen durch Abstraktion der vorgestellten Umgehensweisen mit Metainformation motiviert.

In Kapitel 5 erfolgt dann die Ausarbeitung einer Theorie der Geo-Metainformation. Hier werden die aus der Semiotik benötigten Begriffe definiert und erläutert, welche die Grundlage für eine Analyse der diversen Ansätze bilden. Danach werden mit Hilfe von aus der Semiotik abgeleiteten Setzungen schrittweise die wichtigen Vokabeln definiert, um zu einem semiotischen Verständnis der Metainformation zu gelangen. Diese Definitionen bauen auf den im vorangegangenen Kapitel gewonnenen Erkenntnissen auf und befassen sich mit Begriffen wie Daten, Information, Geoinformation und Umweltinformation. Mit Hilfe dieser Setzungen wird

schließlich das Kernproblem der Arbeit, die Definition von Metainformation und insbesondere ihrer räumlichen Variation durchgeführt.

Abschließend wird in Kapitel 6 diskutiert, für welche Problembereiche die Ergebnisse dieser Arbeit von Bedeutung sind und welchen Einfluss gegenwärtige technische und wissenschaftliche Entwicklungen auf die raumbezogene Metainformationsverarbeitung in Zukunft haben werden. Besondere Beachtung finden hier die sich im Moment weltweit im Aufbau befindlichen und vernetzten Geodaten-Infrastrukturen. Außerdem werden anhand eines praktischen Beispiels, dem Metainformationssystem für Geodaten von Luxemburg, die konzeptionellen Ergebnisse der Arbeit illustriert.

2 Wissenschaftliche Einordnung

Inhalt und Motivation dieses Kapitels ist die Darlegung des wissenschaftlichen Hintergrundes der vorliegenden Arbeit. Eine Studie über geografische Metainformationsverarbeitung sucht die Lösung von Problemen sowohl der raumbezogenen Informatik, als auch der Umweltinformatik, die somit den wissenschaftlichen Rahmen dieser Arbeit stellen. Der Umweltinformatik kommt hierbei eine eher historische Rolle zu, da die Probleme der Metainformationsverarbeitung im Problemkontext des Aufbaus von Umweltinformationssystemen und der Zusammenführung heterogener Umweltdaten erstmalig wahrgenommen wurden. Da zur Lösung der zu behandelnden Probleme Konzepte der Semiotik benutzt werden, bildet eine kurze Abhandlung über die semiotische Wissenschaft die Abrundung des Kapitels.

2.1 Geographical Information Science

Diese Arbeit fügt sich in den Rahmen der *Geographical Information Science (GIScience)*, was sinnvollerweise, wenn auch nicht völlig gleichwertig, als Geoinformatik oder besser Geografische Informationswissenschaft übersetzt werden kann. Die Geoinformatik befasst sich mit dem Wesen und der Funktion der Geoinformation (BARTELME 1994). Mit Geoinformation oder raumbezogener Information wird zum Ausdruck gebracht, dass das Charakteristische dieser Art von Information die Beziehung zu einer Verortung auf der Erdoberfläche ist. Systeme zur Verarbeitung dieser speziellen Art von Information werden Geografische Informationssysteme (GIS) genannt, oder genauer (WORBOYS 1995, S. XI):

Ein Geografisches Informationssystem ist ein computerbasiertes Informationssystem, das die Eingabe, Modellierung, Manipulation, Abfrage, Analyse und Präsentation von raumbezogenen Daten erlaubt.

Das Besondere an einem GIS gegenüber einer herkömmlichen Datenbank oder einem Informationssystem ist die Verwaltung des Raumbezugs, der

grundsätzlich durch eines der zwei Modelle räumlicher Information hergestellt wird. Diese sind zum einen das feldbasierte Modell, welches als Konzept allen Rasterdarstellungen, aber auch anderen überlappungsfreien und vollständigen Mosaizierungen der Erdoberfläche zugrunde liegt, und zum anderen das objektbasierte Modell, welches den Raum als ein Attribut von wohlunterscheidbaren räumlichen Objekten konzeptualisiert (COUCLELIS 1992). In diesem Zusammenhang ist es wichtig zu erwähnen, dass das objektbasierte Modell, das den Vektorimplementationen in GIS zugrunde liegt, nicht gleichzusetzen ist mit objektorientierter Datenbanktechnik oder Programmierung. Es existiert beispielsweise kein Widerspruch darin, einen feldbasierten Modellansatz in einer objektorientierten Programmierweise umzusetzen.

Die Entstehung der Geoinformatik als Wissenschaft wurde durch die Fragen motiviert, die durch den Einsatz dieser Systeme entstanden. Es existierte demnach zunächst die Technologie und erst später befasste man sich mit den konzeptionellen Fragen, die durch den Einsatz der Technologie motiviert wurden. Die Geoinformatik als neuer und interdisziplinärer Fachbereich besitzt Bezüge zu verschiedenen "Mutterwissenschaften", vor allem zur Geografie, der Geodäsie und Kartografie und der Informatik, aber auch zu auf den ersten Blick in diesem Zusammenhang exotisch scheinenden Disziplinen wie der Psychologie. Nach GOODCHILD ET AL. (1999) beschäftigt sich die *GIScience* mit drei Themenbereichen:

- dem **Individuum** als Benutzer von Technologien, Beobachter geographischer Phänomene, Quelle von Konzeptualisierungen und Entscheidungsträger,
- dem **System**, definiert als dem gesamten Komplex digitaler GI-Technologien und der unterstützenden Hardware, Software und den Netzwerkgrundlagen und
- der **Gesellschaft**, mitsamt ihrer Institutionen, ihren Bräuchen, Gemeinschaften, Normen und Standards.

Davon abgeleitet werden drei Aufgabenbereiche, welche die wichtigsten Aspekte der Forschung in der Geoinformatik zusammenfassen (GOODCHILD ET AL., 1999). „Kognitive Modelle des geografischen Raumes“ ist der Titel des

Aufgabenbereiches, der sich mit Fragen des Individuums auseinandersetzt, wobei zu diesem Bereich ebenfalls Fragen der Darstellung und Informationsvermittlung gehören. Erkenntnisse aus der Kognitionswissenschaft, der Umweltpsychologie und der Linguistik sind für diese Art von Forschung von großer Bedeutung. „Computermethoden zur Darstellung geografischer Konzepte“ stehen im Kern der mit dem System befassten Aufgaben und stellen den Bereich dar, welcher in großer Nähe zur Informatik steht. Wirtschaftswissenschaften, Soziologie, Geografie und Politologie tragen zur Lösung von Forschungsfragen zum Thema „Geografien der Informationsgesellschaft“ bei, welches den dritten Forschungsbereich darstellt. Hier werden Fragen behandelt, die sich mit den Veränderungen auseinandersetzen, die durch den Einsatz von GI Systemen innerhalb der Gesellschaft entstehen in Hinsicht auf den Umgang mit Raum.

Die vorliegende Arbeit befasst sich mit der Konzeptualisierung von raumbezogenen Metadaten, kann somit unter dem ersten dieser drei Aufgabenbereiche geführt werden. Das Ziel dieser Art von Forschung ist die Ausarbeitung theoretischer Grundlagen für Geografische Informationssysteme, um eine kohärente Forschungsgemeinschaft zu erzeugen, welche sowohl die Geoinformatik an sich, als auch Geografische Informationssysteme und die dazu beitragenden Felder und Disziplinen weiterbringt (MARK ET AL. 1998).

2.2 Umweltinformatik

Geoinformatik und Umweltinformatik zeichnen sich durch große Überlappungsbereiche aus, da Umweltprobleme eine „problemkonstituierende geografische Dimension“ besitzen (GREVE 1999, S. 33). Ebenso wie die Geoinformatik kann die Umweltinformatik als Teilgebiet der angewandten Informatik betrachtet werden (PAGE & HILTY 1994), besser aber ist es, sie als interdisziplinären Wissenschaftsbereich im Überschneidungsgebiet zwischen Informatik und Anwendungswissenschaften zu sehen, welcher sich allgemein mit der Lösung von informationstechnischen Problemen im Bereich des Umweltschutzes und des Umweltmanagements befasst. Einer der Forschungsschwerpunkte der Umweltinformatik ist die Entwicklung von

Umweltinformationssystemen, die, wenn man sie nicht gleich als erweiterte Geoinformationssysteme auffassen will, integrierte Systeme sind, raumbezogene Informationsverarbeitung in Form von Diensten umfassen. Andere Themenschwerpunkte, die der Umweltinformatik zuzurechnen sind, besitzen ebenfalls zahlreiche Bezüge zur raumbezogenen Informationsverarbeitung. Beispiele hierfür sind die Modellbildung und Simulation sowie das Umweltmonitoring.

Ein Hauptproblem von Umweltinformationssystemen ist die Datenintegration, da Umweltdaten extrem umfangreich und heterogen sind. Da der Aufbau von Umweltinformationssystemen in Deutschland seit etwa 20 Jahren stattfindet, wurde das Problem der Datenintegration schon recht früh erkannt. Dies führte dazu, dass die Auseinandersetzung mit Metadaten in diesem Bereich schon vor einiger Zeit begann. Da Umweltinformationssysteme fast immer auch räumliche Daten verwalten, spielt die Umweltinformatik eine Art Vorläuferrolle in der Diskussion um den Einsatz von Metainformationen. Illustrierend hierfür kann die Entwicklung des Umweltdatenkatalogs (UDK) angeführt werden, dessen anfängliche Entwicklung zu Beginn der 1990er Jahre stattfand und eines der ersten Metainformationssysteme darstellt. Der UDK katalogisiert von Beginn an ebenfalls Geoinformationen (LESSING & WEILAND 1990) und findet inzwischen auch außerhalb des Umweltbereiches seinen Einsatz.

Aus diesem Grund ist es nicht verwunderlich, dass dem Metadatenmanagement, als einem Mittel zur Datenintegration innerhalb der Umweltinformatik, nach wie vor große Aufmerksamkeit gilt. Somit ist eine Theorie der raumbezogenen Metainformationsverarbeitung automatisch auch ein Betrag zur Umweltinformatik.

2.3 Semiotik

Im Mittelpunkt der Semiotik steht der Begriff des Zeichens, beziehungsweise der des Zeichensystems. Als die Wissenschaft der Zeichen befasst sich die Semiotik nicht nur mit den sprachlichen Zeichen, also Problemen die der Linguistik zugeordnet werden, sondern mit allen Zeichensystemen. Die Bereiche einer solchen umfassenden oder allgemeinen Semiotik sind

beispielsweise die Zoosemiotik, die Medizinische Semiotik und die Grafische Semiotik, die sich unter anderem mit den Karten als grafischen Zeichen befasst. Die Disziplin der Semiotik wird selten auch als Semiologie bezeichnet. Die unterschiedlichen Ansätze in der Geschichte dieser Disziplin sind terminologisch nicht konsistent. In dieser Studie wird der Begriff Semiotik angelehnt an die Arbeiten von ECO (1987) verwendet.

Die Beziehung der *GIScience* zur Semiotik besteht in erster Linie durch die Interpretation von Karten als Zeichen, womit sich die Grafische Semiologie befasst (BERTIN 1974). Karten, sowohl in digitaler als auch in analoger Form, sind Informationsträger, die mittels eines Kodes verschlüsselt sind.

Zeichensysteme wie die gesprochene Sprache können als lineare Zeichensysteme bezeichnet werden, da sie nur über zwei Variablen verfügen und zwar über die Variation der Töne und die Zeit (BERTIN 1974). Schriftsprache ist eine Transkription dieses Zeichensystems, welche lediglich die Speicherung dieser Zeichen vornimmt und somit ebenfalls linear ist. Räumliche Zeichensysteme verfügen über spezielle Vorteile gegenüber linearen Systemen wie der Sprache weil sie Beziehungen zwischen drei Variablen nämlich den beiden Dimensionen der Ebene und der „Variation der Flecken“ (BERTIN 1974, S. 11) vermitteln können. Grundlegend sind Karten spezielle Darstellungsweisen für Informationen, die auch in Textform abgebildet werden können. Die drei Variablen der grafischen Darstellung sind hierbei, wie bereits angesprochen, die beiden Dimensionen der Ebene und die inhaltstragende Variable. Diese Informationen können beispielsweise ebenfalls in Tabellenform dargestellt werden, wobei dann sequenziell immer nur eine Beziehung der drei Variablen aufgenommen und verarbeitet werden kann. In der Kartendarstellung wird die dritte Variable durch grafische Symbole oder Farbvariationen transkribiert und zu allen anderen Teilinformationen mittels des Kartenbildes in Beziehung gesetzt. Hierdurch können mehr Informationen in kürzerer Zeit an den Informationsempfänger übertragen werden, wenn dieser die korrekte Entschlüsselungsmethode des Zeichensystems beherrscht. Grafische Zeichensysteme sind anderen Systemen somit im Hinblick auf die Informationsvermittlung überlegen.

2.4 Zusammenfassung

Theoriebildung, die sich mit raumbezogener Metainformation befasst, ist ein Beitrag sowohl zur Geografischen Informationswissenschaft, da raumbezogene Metainformation selbstverständlich auch Geoinformation ist, als auch zur Umweltinformatik. Die Gründe hierfür liegen darin, dass Umweltinformationen erstens im Allgemeinen einen Raumbezug besitzen und dass zweitens die Metainformationsverarbeitung ein wichtiges Thema in der Umweltinformatik darstellt. Die Beziehung des Arbeitsthemas zur Semiotik liegt nicht direkt auf der Hand, ist aber über die Grafische Semiologie bereits vorhanden. Die direkte Verknüpfung zwischen Semiotik und Metainformationsverarbeitung stellt ein Ziel dieser Arbeit dar. Sinn dieser Verbindung ist die Konstruktion einer stabilen konzeptionellen Grundlage für die Geo-Metainformationsverarbeitung und eine größere Klarheit der assoziierten Begriffe.

Die vorliegende Arbeit ist hiermit in ihren wissenschaftlichen Rahmen gestellt, im nächsten Kapitel wird der konzeptionelle Referenzrahmen und die Vorgehensweise diskutiert.

3 Referenzrahmen und Methodik

Dieses Kapitel dient zur Erläuterung der methodischen Vorgehensweise und zur Fokussierung des Problembereiches. Als Einführung in die dieser Arbeit zugrundeliegende Vorgehensweise ist es hilfreich, einige allgemeine Gedanken über die Natur von Wissenschaft darzulegen und sich mit ausgewählten epistemologischen Fragen auseinander zu setzen. Hierdurch wird klargestellt, was das Produkt dieser Arbeit ist und wie es in Beziehung zur Semiotik steht. Die Darlegungen der folgenden beiden Abschnitte sind angelehnt an die Ausführungen von ABLER, ADAMS & GOULD (1971) und dienen zur Eingrenzung des Problembereiches der vorliegenden Arbeit. Die Erkenntnisse dieser Autoren werden deshalb hier genutzt, da sie ein konsistentes, einfaches und stark auf die Geografie bezogenes Wahrnehmungsmodell geschaffen haben, das sich auf sehr sinnvolle Weise zur Darlegung der angewandten Vorgehensweise adaptieren lässt. Eine Abhandlung über den Theorie- und Modellbegriff vervollständigt diese wissenschaftstheoretische Definition des Referenzrahmens. Durch Anwendung der Konzepte des *OpenGIS Consortiums* wird der Referenzrahmen auf einer eher ingenieurstechnischen Ebene komplettiert. Am Ende dieses Kapitels wird die verwendete Methodik erklärt, die sich auf die Anwendung semiotischer Konzepte auf Phänomene raumbezogener Informationsverarbeitung stützt.

3.1 Aufbau und Natur von Wissenschaft

Der praktische Erfolg einer Wissenschaft misst sich im Allgemeinen an ihrer Fähigkeit, Lösungen für Probleme zu liefern. Eine nach dieser Definition erfolgreiche und somit funktionierende Wissenschaft setzt sich zu unterschiedlich großen Teilen aus Praktikern, Methodikern, Theoretikern und Philosophen zusammen. Abb. 1. veranschaulicht diesen Zusammenhang in Form einer Pyramide.

Die Form der Pyramide und die Verteilung der vier Gruppen deutet bereits darauf hin, dass Praktiker, also Menschen, die konkrete Probleme zu lösen

haben, die größte Anzahl an Personen in der Pyramide einnehmen sollten. Praktische Problemlösung alleine kann allerdings innerhalb einer Wissenschaft nicht existieren. Zu ihrer Weiterentwicklung bedarf es Menschen mit anderen Interessen innerhalb des Wissenschaftsgebäudes. Methodiker sind Personen, die sich darüber Gedanken machen, auf welche Art und Weise man Probleme lösen sollte. Zur Arbeit des Methodikers ist ein gewisses Maß an Abstraktion nötig. Hierdurch kann er von konkreten Problemen zu allgemeinen theoretischen Aussagen über deren Lösungsmöglichkeiten gelangen. Die Methodik einer Wissenschaft stellt das Rüstzeug zur Verfügung, mit der die Praktiker Probleme lösen können. Umgekehrt motivieren Probleme der Praxis die Weiterentwicklung der Methodik.

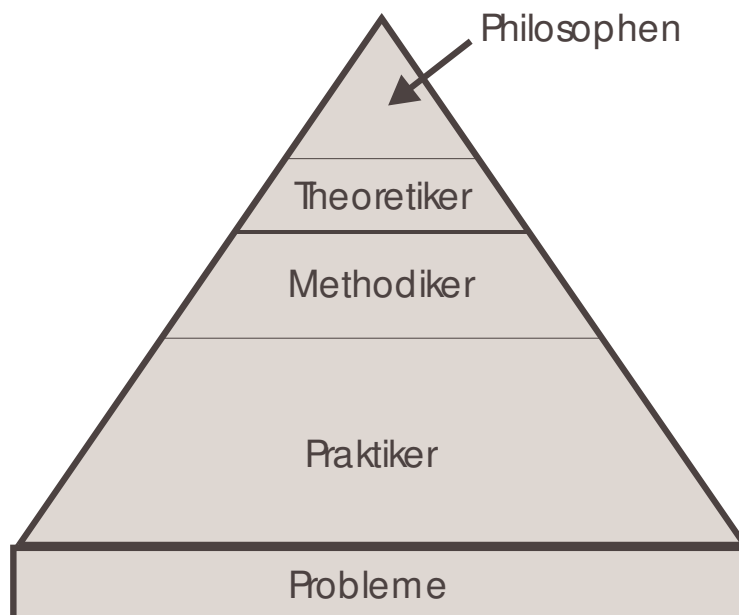


Abbildung 1. Struktur einer Wissenschaft (Quelle: nach Abler, Adams & Gould., 1971)

Eine noch kleinere Gruppe von Personen, die Theoretiker, befasst sich mit Fragen, die noch weiter von der konkreten Problemlösung entfernt sind. Sie denken darüber nach, auf welche Art und Weise Angehörige ihrer Disziplin über das, was sie tun, denken. Sie befassen sich also auf einer noch höheren Abstraktionsebene mit ihrer Wissenschaftsdisziplin. Die praktische Lösung von Problemen ist für Theoretiker zweitrangig, wengleich die Theoriebildung den geistigen Unterbau darstellt, auf den Methodiker und Praktiker aufbauen. Am weitesten entfernt von unmittelbaren Problemen stehen die Philosophen, die sich ganz allgemein damit befassen, wie Menschen denken und über was

sie am besten nachdenken sollten. Philosophen arbeiten an Konzepten, die über den Kontext einer einzelnen Wissenschaftsdisziplin hinausgehen und folglich disziplinübergreifend sind.

Die Entwicklung eines allgemeinen Modells der raumbezogenen Metainformationsverarbeitung ist ein Beitrag zur Theoriebildung der Geografischen Informationswissenschaft und somit im Aufgabengebiet eines Theoretikers dieser Disziplin anzusiedeln. Zur Lösung dieser Aufgabe sollen Konzepte der Semiotik herangezogen werden. Semiotik, die Lehre der Zeichensysteme, ist außerdem Teil der Philosophie, beziehungsweise ist die allgemeine Semiotik eine Art der Sprachphilosophie (ECO 1985). Ihre Wurzeln hat die Semiotik in der Linguistik. Allgemeine Semiotik befasst sich mit dem Begriff des Zeichens als Informationsträger und der dazugehörigen Kommunikationsprozesse. Dies umfasst die sprachliche Kommunikation, aber auch alle anderen Arten von Zeichen, wie beispielsweise die Symptome, die ein Arzt zum Diagnostizieren einer Krankheit benutzt, oder alle Arten von grafischen Zeichen wie Diagramme oder Karten.

Innerhalb der Semiotik kann zwischen allgemeiner und spezieller Semiotik unterschieden werden. Die allgemeine Semiotik hat den eben definierten, über die Linguistik hinausgehenden Umfang. Eine spezielle Semiotik befasst sich mit der „Grammatik“ eines bestimmten Zeichensystems. Sie versucht die Beschreibung eines gegebenen Feldes kommunikativer Phänomene. Eine spezielle Semiotik muss sich mit erkenntnistheoretischen Fragestellungen befassen, um ihr Objekt zu definieren und abzugrenzen (ECO 1985). Raumbezogene Informations- und Metainformationsverarbeitung stellt das Feld kommunikativer Phänomene dar, mit dem sich diese Arbeit befasst. Folglich ist die Anwendung semiotischer Konzepte in diesem Feld die Entwicklung einer speziellen Semiotik.

3.2 Abgrenzung des Problembereiches

Anhand eines von ABLER, ADAMS & GOULD (1971) erarbeiteten Referenzrahmens zur Ordnung von Erfahrungen wird nun der Problembereich dieser Arbeit erläutert.

Das epistemologische Schema, welches in Abb. 2. dargestellt ist, geht davon aus, dass Wahrnehmung und menschliches Denken allgemein, aber insbesondere wissenschaftliches Arbeiten, sich im Kern mit der Ordnung von Ereignissen befasst.

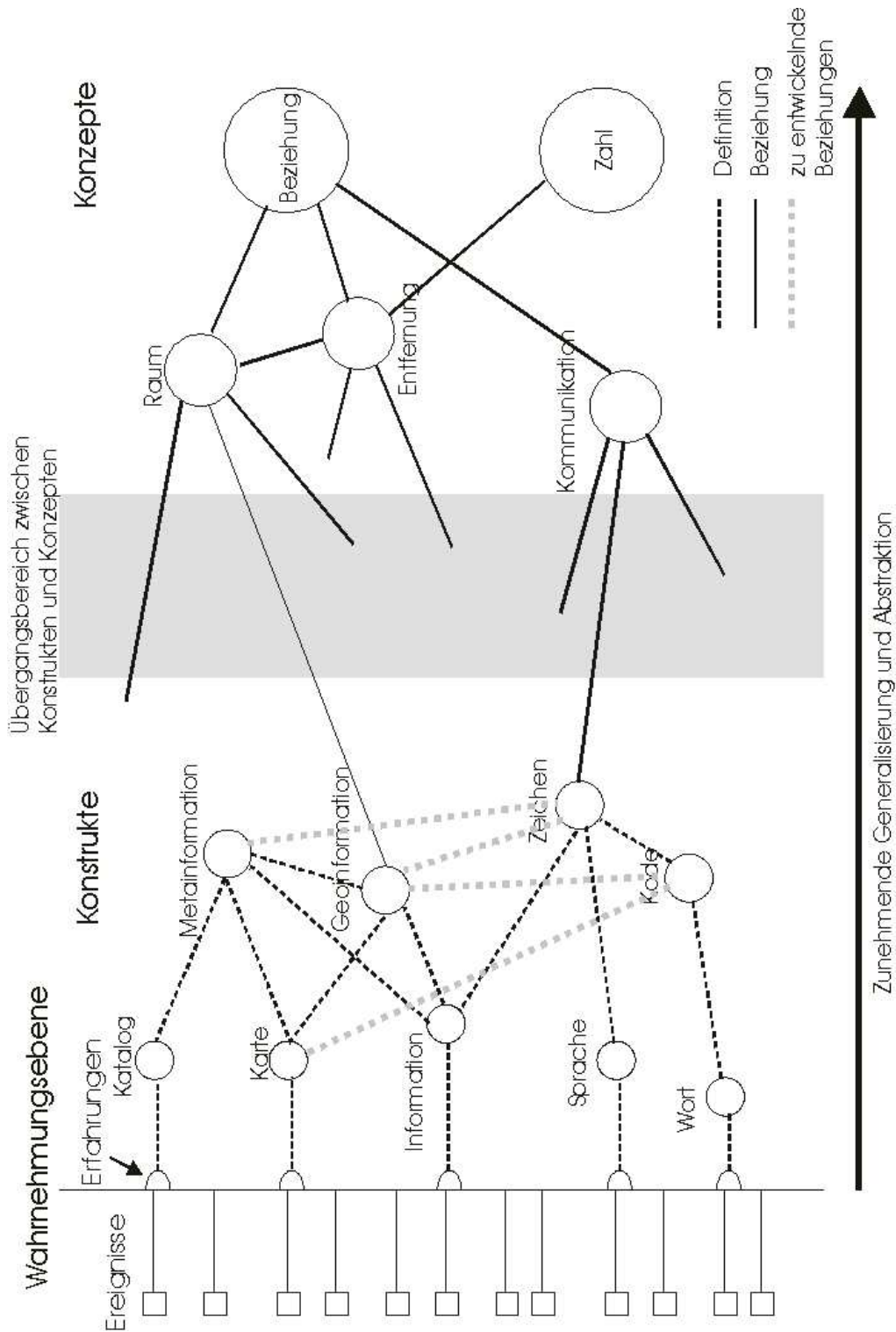


Abbildung 2. Ordnung von Erfahrungen (Quelle: nach Abler, Adams & Gould., 1971, verändert)

Eine Unmenge von Ereignissen geschieht ständig um uns herum, aber nur ein Bruchteil dieser Ereignisse werden auch als solche erkannt und erscheinen auf unserer Wahrnehmungsebene. Wahrgenommene Ereignisse wollen wir als Erfahrungen bezeichnen. Wir beginnen, Erfahrungen zu ordnen, indem wir Konstrukte bilden und den Erfahrungen diese zuordnen. Konstrukte sind bestimmte Ideen über die Erfahrung, die eine erste Ordnung erlauben oder sogar notwendig sind, um aus einem Ereignis überhaupt eine Erfahrung zu machen. Konstrukte existieren auf verschiedenen Stufen der Abstraktion, beispielsweise ist „Rose“ ein recht spezielles Konstrukt, das dem allgemeineren „Blume“ und dieses wiederum dem noch allgemeineren Konstrukt „Pflanze“ zugeordnet werden kann. Konstrukte sind zwar Verallgemeinerungen, man kann ihnen allerdings empirischen Inhalt zuweisen. Auf noch höherem Abstraktionsgrad erschaffen wir Konzepte, welche keinen empirischen Inhalt mehr aufweisen. Konzepte werden aufgebaut auf einer Vielzahl von Wahrnehmungen und Konstrukten. Die meisten Konstrukte und Konzepte werden allerdings nicht von uns selbst durch den Vorgang der Abstraktion erzeugt, sondern von Anderen an uns weitergegeben. Der Übergang zwischen Konstrukten und Konzepten ist fließend. Es existiert eine Unmenge von Konzepten wie beispielsweise Entfernung, Form, Geschwindigkeit und Menge. Alle Konzepte können auf die Megakonzepte Zahl und Beziehung reduziert werden. Diese beiden Konzepte können universal benutzt werden.

In Abb. 2. zeigen gestrichelte Linien Definitionen an. Definitionen sind Beziehungen zwischen Erfahrungen und Konzepten. Durchgezogene Linien stehen für andere Arten von Beziehungen, die zwischen zwei Konstrukten oder einem Konstrukt und einem Konzept existieren können. Die in Abb. 2. gezeigten Erfahrungen und Konzepte zeigen Aspekte der vorliegenden Arbeit. Im oberen Teil sind die Konstrukte aus dem Bereich der Geografischen Informationswissenschaft dargestellt, im unteren aus dem Bereich der Semiotik. Die grauen gepunkteten Linien deuten die Verbindungen an, deren Herstellung Hauptaufgabe der vorliegenden Arbeit darstellt.

3.3 Theorie- und Modellbegriff

Durch Benutzung semiotischer Begriffe und Denkweisen und der Anwendung dieser Konzepte auf die Metainformationsverarbeitung wird eine Teiltheorie der Semiotik definiert. Ob man ein solches Gedankengebäude Theorie oder Modell nennt, ist eine Diskussion die der Philosophie oder genauer der Erkenntnistheorie zuzuordnen ist. Beispielsweise definieren KRIZ ET AL. (1990) Theorie als System von Aussagen. Aus Sicht eines am Modellbegriff orientierten systematischen Neopragmatismus ist ein Modell ganz allgemein ein Erkenntnisgebilde und umfasst somit auch die Theorie (STACHOWIAK 1983). Nach ABLER, ADAMS & GOULD (1971) sind Theorien Strukturen, die sich aus wissenschaftlichen Gesetzen oder gesetzesähnlichen Aussagen und deren Verbindungen zusammensetzen. Das hervorstechende Merkmal einer Theorie nach dieser Definition ist demnach die Verbindung eines Systems von Aussagen, die einen höheren Abstraktionsgrad liefert. Die selben Autoren sehen zwei Möglichkeiten den Modellbegriff zu definieren. Im engeren Sinne ist ein Modell ein struktureller Isomorphismus zwischen zwei Theorien, wobei ein Isomorphismus eine eins-zu-eins Beziehung darstellt. Wenn eine eins-zu-eins Beziehung zwischen allen Elementen der einen Theorie zu allen Elementen der zweiten Theorie existiert, sind also die Voraussetzungen für diesen eng gefassten Modellbegriff erfüllt. Die zweite Definition des Modellbegriffs nach ABLER, ADAMS & GOULD (1971) ist die weiter gefasste, wie sie allgemein in den Sozialwissenschaften benutzt wird. Hier repräsentiert ein Modell entweder eine Hypothese über eine einfache oder komplexe Beziehung oder über die Struktur von Beziehungen. Um eine letzte Sichtweise auf den Modellbegriff hier anzuführen, soll zusätzlich die Definition von WORBOYS (1995) dargestellt werden. Hier wird der Begriff Modell ganz allgemein als strukturerhaltende Abbildung verstanden.

Um die Diskussion über den Theorie- und Modellbegriff abzuschließen, lässt sich zusammenfassend sagen, dass die Umgehensweise mit den beiden Begriffen sehr uneinheitlich ist. Es gibt aber gewisse Grundtendenzen, die den genannten Autoren gemeinsam sind. Charakteristisch für eine Theorie ist, dass sie aus einer Mehrzahl von Aussagen besteht, die durch eine gemeinsame Idee oder Abstraktion verknüpft sind. Die wichtigste Charakteristik eines Modells

ist die Abbildung zwischen zwei Konstrukten, seien dies nun Theorien oder realweltliche Phänomene und ihre mathematischen Repräsentationen. Der in dieser Arbeit verfolgte Ansatz benutzt als Grundlage die semiotische Theorie, welche im Übrigen alle Merkmale einer echten Theorie im Sinne von ABLER, ADAMS & GOULD (1971) erfüllt. Diese wird verknüpft mit Konzepten und Phänomenen der raumbezogenen Informationsverarbeitung, es wird also eine Abbildung zwischen diesen und den semiotischen Konzepten erstellt. Folglich kann man von einem Modell der raumbezogenen Metainformationsverarbeitung sprechen, welche zur Bildung einer Teiltheorie der Semiotik, einer speziellen Semiotik, führt.

3.4 Beziehung zum OpenGIS-Kontext

Wie bereits ausgeführt stellen die Konzepte des *OpenGIS Consortium* (OGC) einen ingenieurwissenschaftlichen Referenzrahmen dieser Arbeit dar. Um die Konzepte des OGC einzuführen, wird zuerst das Interoperabilitätsproblem erläutert, welches Grundlage und Motivation der Arbeit des OGC darstellt.

Die Anzahl der Dokumente des OGC ist inzwischen sehr umfangreich geworden, so dass sich diese Arbeit auf die Vorstellung der für die raumbezogenen Metainformationsverarbeitung wichtigen Dokumente beschränkt. Dies sind die Dokumente, welche sich mit den abstrakten Konzepten auf oberster Ebene befassen sowie einige Dokumente, die Metadatenverarbeitung und Geodatenkataloge behandeln.

3.4.1 Das Interoperabilitätsproblem

Eines der derzeitigen Hauptprobleme im Bereich GIS ist das Interoperabilitätsproblem, welches sich mit dem Informationsaustausch zwischen GISystemen befasst. Probleme dieser Art werden beispielsweise durch unterschiedliche Arten der (Geometrie-)Datenmodellierung oder durch abweichende Belegung von Begriffen in verschiedenen Anwenderdomänen verursacht. Interoperabilitätsprobleme können mit Hilfe von Metainformation gelöst, beziehungsweise deren Lösung unterstützt werden. Grundsätzlich kann zwischen syntaktischen und semantischen Interoperabilitätsproblemen unterschieden werden. BISHR (1998) unterscheidet semantische, schematische

und syntaktische Probleme. Eine ähnliche Unterscheidung wird von DENZER & GÜTTLER (1994) im Umweltbereich benutzt und zwar zwischen semantischer, syntaktischer und struktureller Heterogenität, wobei letztere der schematischen von BISHR (1998) entspricht. Zusätzlich wird hier noch die dynamische Heterogenität angeführt, welche allerdings nur dem zeitlichen Aspekt der anderen beiden Heterogenitätsaspekte entspricht. Schematische/Strukturelle Heterogenität umfasst Aspekte semantischer und syntaktischer Art. Es ist somit abstrahierend die Aussage möglich, dass es nur zwei Arten der Heterogenität von Datenquellen und somit Interoperabilitätsproblemen gibt. Diese Sichtweise wird von der *International Standards Organisation* geteilt (ISO 2001). Die beiden Arten von Interoperabilitätsproblemen sind:

1. Syntaktische Interoperabilitätsprobleme, die durch verschiedene Technik der verwendeten Systeme entstehen (objektorientierte versus relationale Datenbanken, verschiedene Datenformate, ...) und
2. Semantische Interoperabilitätsprobleme, die durch verschiedene Sichten auf die durch die Daten repräsentierten Phänomene entstehen.

Viele syntaktische Interoperabilitätsprobleme können mit Hilfe von Standardisierungen gelöst werden. Beispielsweise wird die Übersetzung zwischen verschiedenen Raumbezugssystemen durch die allgemeine Verwendung der von der *European Petroleum Survey Group*¹ (EPSG) zur Verfügung gestellten geodätischen Parameter gelöst.

Semantische Interoperabilitätsprobleme entstehen durch die unterschiedlichen Gedankenmodelle verschiedener Benutzergruppen (BISHR 1998). Um diese Probleme zu überwinden, ist es hilfreich, den Entstehungskontext mit Hilfe von Metadaten zu dokumentieren und somit für potenzielle Nutzer verständlich zu machen, oder mittels inhaltlicher Standardisierung zu homogenisieren (GREVE 1998).

Metainformationssysteme, die Geo-Metadaten verarbeiten, speichern und präsentieren, dienen üblicherweise der Integration von Geodaten aus verschiedenen Quellen. Da Metainformationssysteme aber wiederum selbst

¹ sieht <http://www.epsg.org>

Informationssysteme sind, haben sie ebenfalls mit Interoperabilitätsproblemen zu kämpfen. Es besteht somit ein Interoperabilitätsproblem auf Metadatenebene, welches sich gleichfalls in eine syntaktische und eine semantische Komponente unterteilen lässt.

Das internationale *OpenGIS Consortium* ist ein Zusammenschluss einer großen Zahl von Institutionen, die im Bereich der raumbezogenen Informationsverarbeitung tätig sind. Firmen verschiedener Größe, Verwaltungs- und Forschungseinrichtungen aus einer Vielzahl von Nationen, vor allem aus Nordamerika, aber auch aus Europa und Asien, arbeiten in diesem Forum zusammen. Das OGC hat sich zum Ziel gesetzt, die Interoperabilität in der geographischen Informationsverarbeitung und die Integration von raumbezogener Informationstechnologie in Standard IT-Verfahren voranzutreiben. Zu diesem Zweck verfasst das OGC mittels eines festgelegten Konsensprozesses Spezifikationen, die definierte Schnittstellen beschreiben. Mittels dieser Schnittstellen ist es möglich, Systeme anzusprechen und in Anwendungen raumbezogener Informationsverarbeitung zu integrieren. Der Erfolg des OGC bedeutet letztendlich den Abschied von den bislang verbreiteten monolithischen Geografischen Informationssystemen. Softwareprodukte, die diese Schnittstellen implementieren, können von Seiten des OGC einem Test unterzogen werden, um auf ihre Konformität geprüft zu werden. Dies erlaubt den Herstellern ein offizielles OGC Konformitätsabzeichen zu führen.

Das OGC befasst sich bislang schwerpunktmäßig mit der Definition von Schnittstellen, die syntaktische Interoperabilitätsprobleme lösen. Die Arbeit des OGC an der Lösung semantischer Interoperabilitätsprobleme hat bislang noch keine umsetzbaren Ergebnisse gebracht. Innerhalb einer speziellen Interessengruppe des OGC werden Probleme dieser Art allerdings diskutiert².

Das *Technical Committee* (TC) 211³ der *International Standards Organisation* (ISO) befasst sich ebenfalls mit Standardisierung im Bereich geographischer Informationsverarbeitung. Die Spezifikationen von ISO/TC 211 sind im Gegensatz zu denen des OGC allerdings auf abstrakterer Ebene angesiedelt.

² und zwar innerhalb der *Information Communities and Semantics SIG*

³ Informationen zu ISO/TC 211 sind zu finden unter <http://www.isotc211.org/>

Seit Gründung des ISO/TC 211 1994 besteht eine *Class A Liason* zwischen OGC und ISO/TC 211, welche die Zusammenarbeit der beiden Gremien regelt. Seit 1997 wurde diese Zusammenarbeit noch verstärkt und der Beschluss gefasst, die Spezifikationen zu harmonisieren (KUHNS 1997). Diese Zusammenarbeit ist auch wegen der unterschiedlichen Ausrichtung der beiden Organisationen von großer Bedeutung. Das OGC arbeitet technologiebezogen und zielt darauf ab, in möglichst kurzer Zeit technische Spezifikationen zu erarbeiten. Diese sollen Herstellern von GI-Software ermöglichen, interoperable Produkte zu erstellen. Die Entwicklung von ISO-Standards benötigt im Allgemeinen eine längere Laufzeit, jedoch finden sie eine bedeutend größere Beachtung und Verbreitung. Inzwischen finden regelmäßig gemeinsame Treffen der beiden Standardisierungsgremien statt, welche zu einem regen Austausch in beide Richtungen führte. Ergebnis dieses Austausches ist, dass zum einen einige ISO-Standards inzwischen zu abstrakten OGC-Spezifikationen wurden und andererseits eine Zahl von OGC Implementationsspezifikationen den Status eines internationalen ISO-Standards erhielten.

Das OGC entwickelte als Grundlage seiner Spezifikationen ein abstraktes Modell der raumbezogenen Informationsverarbeitung. Dieses Modell, das auch den Bereich Metainformationsverarbeitung mit einschließt, wird mit großer Wahrscheinlichkeit die Zukunft der *GIScience*, zumindest im angewandten Bereich stark beeinflussend wirken. Das Zusammenwirken des OGC-Modells und des in dieser Arbeit entwickelten Denkmodells ist deshalb sehr vorteilhaft. Das abstrakte OGC-Modell stellt einen ingenieurtechnischen Referenzrahmen für die vorliegende Studie dar. Die in dieser Arbeit entwickelte Theorie versteht sich als Präzisierung und Erweiterung des OGC-Modells und steht in keiner Weise im Widerspruch dazu. Um dies zu erreichen, werden nun die für die raumbezogene Metainformationsverarbeitung relevanten Teile der OGC-Spezifikationen vorgestellt und im Weiteren, soweit anwendbar, die Terminologie des OGC verwendet.

3.4.2 Relevante OGC-Spezifikationen

Der *OpenGIS Guide* (BUEHLER & MCKEE 1998) stellt die Einführung in die Sichtweise des OGC auf den gesamten Bereich der interoperablen raumbezogenen Informationsverarbeitung dar. In diesem Dokument findet man eine grundlegende Einführung in die Motivation und den Ansatz, den das OGC verfolgt. Ziel des OGC ist zum einen die Integration raumbezogener Daten sowie raumbezogene Daten verarbeitende Ressourcen in die allgemeine Informationsverarbeitung. Zum anderen soll die verbreitete Nutzung von interoperabler, kommerzieller GI-Software als Teil der globalen Informationsinfrastruktur angestoßen werden (BUEHLER & MCKEE 1998). Hierzu werden innerhalb des OGC Spezifikationen entwickelt, die diese Ziele ermöglichen sollen. Das OGC verfasst Spezifikationen auf drei Ebenen, zum einen auf der Ebene des *Essential Model*, welches die Struktur des modellierten Ausschnittes der realen Welt beschreibt und das in der *Abstract Specification* beschrieben wird. Dieses abstrakte Modell gibt den konzeptionellen Rahmen vor, der es erlaubt Technologie herzustellen, welche die folgenden Kommunikationsbarrieren zu überwinden hat:

- Zwischen Systemen: verschiedene, den Raumbezug verarbeitende Softwaresysteme, die unterschiedliche Datenformate benutzen.
- Zwischen verteilten Rechenplattformen⁴: ähnlichen oder unähnlichen Raumbezug verarbeitende Softwaresysteme, die auf verschiedenen Plattformen arbeiten.
- Zwischen Informationsgemeinschaften: Gruppen von Nutzern von GI-Software, die geografische Objekte auf ähnliche oder verschiedene Weise definieren.

Das *Essential Model* muss sowohl allgemein als auch detailliert genug sein, um die genannten drei Ziele erreichbar zu machen. Hierzu wird die gesamte Bandbreite der raumbezogenen Informationsverarbeitung wie Darstellung, Recherche, Zugriff und Verarbeitung abstrakt beschrieben.

Die *Abstract Specification* umfasst außerdem das *Abstract Model*, welches die zweite Spezifikationsebene darstellt. Die *Abstract Specification* ist in eine

⁴ In OGC-Terminologie entspricht das den *Distributed Computing Platforms* (DCPs)

Anzahl von *Topics* unterteilt, die sich mit speziellen Teilaspekten raumbezogener Informationsverarbeitung befassen. In diesen *Topics* werden die Teilaspekte auf abstrakter Ebene spezifiziert und die Anforderungen an die das *Topic* implementierende Software erläutert. Dies umfasst die Aufgaben und Funktionalitäten der Software, aber auch Objektmodelle, die den Anwendungsbereich abdecken. Schnittstellen und Objektklassen sind ebenfalls bereits Gegenstand des *Abstract Models*. Sowohl das *Essential Model* als auch das *Abstract Model* bewegen sich auf implementationsunabhängiger Ebene im Gegensatz zum *Specification Model*, das ein allgemeines Modell der Software und ihrer Schnittstellen beschreibt und in den *Implementation Specifications* beschrieben wird. Das *Specification Model* stellt die dritte Spezifikationsebene dar.

Auf Ebene der *Abstract Specification* sind für die Metainformationsverarbeitung folgende Teile relevant:

- *Topic 5 Features* (OGC 1999a)
- *Topic 9 Quality* (OGC 1999b)
- *Topic 11 Metadata* (OGC 1999c)
- *Topic 13 Catalogs* (OGC 1999d)

Da Metadaten nicht nur zur Beschreibung von Datensätzen, sondern auch von Diensten eingesetzt werden können und diese Szenarien zunehmend an Bedeutung gewinnen, ist es zudem notwendig, sich mit der Definition und dem Verhalten von Diensten im Sinne des OGC auseinanderzusetzen. Das *Topic 12* trägt den Namen *Service Architecture* (OGC 2001a), ist aber in der aktuellen Version 4.2 kein eigenständiges Dokument mehr, sondern enthält die *ISO Specification 19119 Geographic Information – Services* (ISO 2001). Dieses Dokument stellt folglich die Grundlage der Sichtweise und die Definition von Geo-Diensten in dieser Studie dar.

Zudem ist von den bislang verfügbaren *Implementation Specifications* die *Catalog Interface Implementation Specification* (OGC 1999e), die eine Implementationsspezifikation für das *Topic 13* darstellt, von Relevanz für das Thema. Die genannten Spezifikationen stellen zusammen mit dem *OpenGIS Guide* (BUEHLER & MCKEE 1998) die Grundlage des nächsten Abschnittes dar und werden dort näher erklärt.

3.4.3 Metainformation und *OpenGIS*

Das *OpenGIS Essential Model* analysiert die Problematik der raumbezogenen Informationsverarbeitung mit Hilfe von neun Abstraktionsebenen (Abb. 3.):

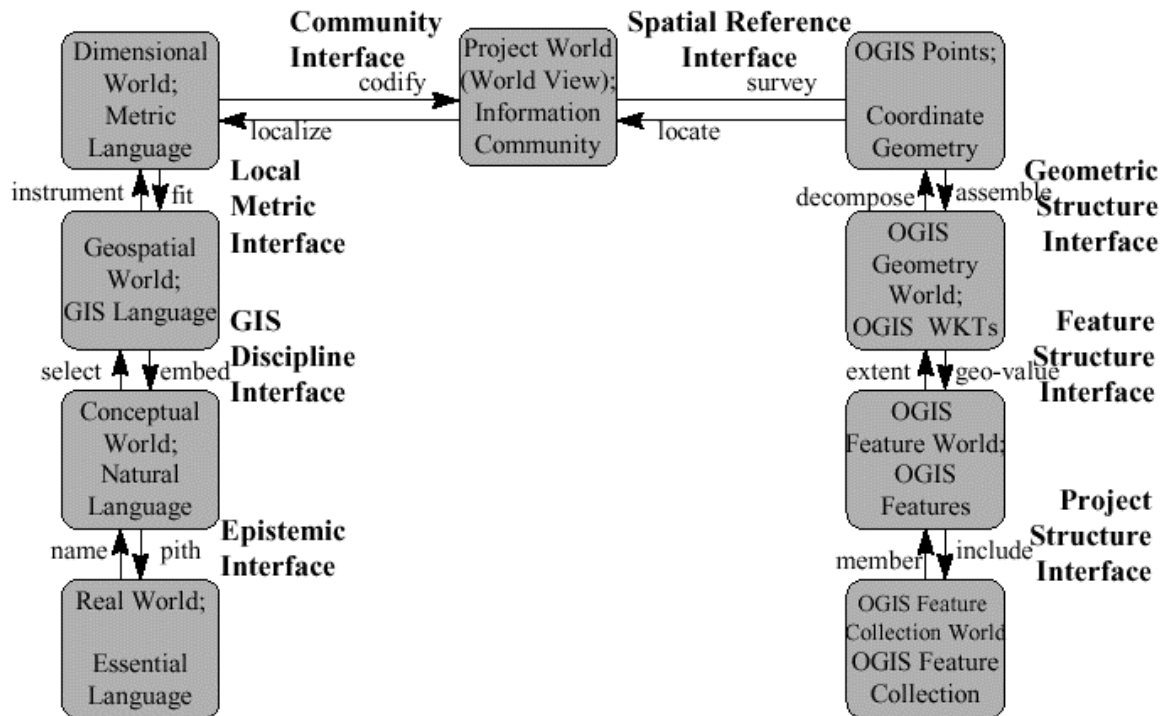


Abbildung 3. Neun Ebenen der Abstraktion (Quelle: OGC 1999a)

1. Reale Welt
2. Welt der Konzepte: entspricht den Dingen, die erkannt und benannt wurden
3. GI-Welt; Ausschnitt aus der Welt der Konzepte, der sich mit raumbezogener Information befasst
4. Dimensionale Welt: die GI-Welt, nachdem sie gemessen wurde
5. Projektwelt: Ausschnitt aus der dimensionalen Welt, der für ein spezifisches Projekt relevant ist
6. *OpenGIS* Punkte: die Definition von Punkten, so dass sie für alle Softwaresysteme nutzbar sind
7. *OpenGIS* Geometrie: die Art, in der Geometrien mit Hilfe von Punkten konstruiert werden
8. *OpenGIS Feature*: die Art, in der *Features* aus *OpenGIS* Geometrien, Attributen und einem Raumbezugssystem konstruiert werden

9. OpenGIS Feature Collection: Sammlungen von Features.

Da der Begriff *Feature* sich inzwischen im deutschen Sprachbereich weitgehend eingebürgert hat und es keine eindeutige Übersetzung gibt, wird im Weiteren auf eine Übersetzung verzichtet. Ein raumbezogenes Phänomen wird also durch Modellierung zu einem *Feature*, welches die grundlegende Einheit eines raumbezogenen Objektes darstellt. Ein *Feature* kann rekursiv definiert werden, so dass sowohl ein Straßenabschnitt zwischen zwei Punkten, als auch ein vollständiges Gewässernetz oder ein georeferenziertes Satellitenbild ein *Feature* darstellen kann. Gruppen von *Features* können zu *Feature Collections* zusammengefasst werden (Abb. 4.). *Feature Collections* sind im Sinne des OGC Spezialisierungen von *Features*. Da *Feature Collections* aus *Features* bestehen, diese aber wiederum *Features* sein können ergibt sich eine rekursive Definition, die es erlaubt, beliebig viele Ebenen von *Feature Collections* zu definieren.

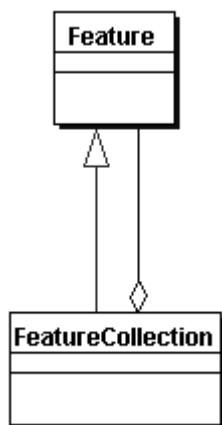


Abbildung 4. Beziehung zwischen Feature und Feature Collection (Quelle: nach OGC 1999a)

Features werden aus drei Elementen aufgebaut:

- Geometrie,
- Semantische Eigenschaften und
- Metadaten

Metadaten werden im *Essential Model* definiert als „[...]andere (als semantische Eigenschaften) Informationen, die benötigt werden könnten, um das Phänomen im Zusammenhang der Anwendungsumgebung oder Anwendergemeinschaft positionieren zu können. [...] Metadaten sind einfach eine Teilmenge der Eigenschaften eines *Features* (oder üblicherweise einer

Feature Collection). [...] Die Komplexität der Metadaten kann angepasst werden, um die Bedürfnisse der Anwendung zu erfüllen. [...]“ (übersetzt aus BUEHLER & MCKEE 1998, S. 42). Es wird also keine ausdrückliche Unterscheidung zwischen den allgemeinen Eigenschaften eines *Features* und den Metadaten vorgenommen, sie beschreiben den Anwendungskontext eines *Features* oder einer *Feature Collection*.

Topic 11 der *Abstract Specification* (OGC 1999c) befasst sich mit Metadaten. Hier wird erklärt, wie Metadatenverarbeitung im OGC-Kontext geplant ist. Metadaten können entweder mit einem *Feature* oder einer *Feature Collection* assoziiert sein. Metadaten, die mit einer *Feature Collection* assoziiert sind, können sich entweder auf die gesamte *Feature Collection* beziehen, auf eine Teilmenge der Elemente oder auf jedes einzelne Element. Metadaten, die mit einem *Feature* assoziiert sind, können sich auf das *Feature* als Ganzes oder auf bestimmte relevante Teilkomponenten beziehen. In diesem Dokument wird wiederholt auf den ISO Standard 19115 *Geographic Information – Metadata* (ISO 2000) verwiesen, welcher die Definition standardisierter Metadatenelemente und –einheiten liefert. Das OGC wird hierzu keine eigenen Spezifikationen veröffentlichen, sondern auf das genannte Dokument zurückgreifen. Außerdem werden anhand von zehn Szenarien die möglichen Anwendungsfälle für Metadaten erläutert und die notwendigen Klassen und Beziehungen für Metadaten dargestellt.

Abstract Specification Topic 9 – Quality (OGC 1999b) beschreibt ein konzeptionelles Modell des Umgangs mit einer speziellen Art von Metadaten, jenen, die sich mit Datenqualität befassen. Die beschriebenen Metadaten beschreiben die Genauigkeit von Geometrien, Bildern und *Coverages*, als auch die Wertegenauigkeit von nichtgeometrischen Attributen. *Coverages* sind in der Terminologie des OGC *Features*, deren Geodaten feldbasiert modelliert wurden. Wie schon in *Topic 11* erläutert, soll jedem *Feature* ein Metadatensatz zugeordnet werden, welcher dann Angaben zur Genauigkeit des *Features* gibt. Die Art der zur Verfügung gestellten Angaben ist dabei abhängig vom Typ des *Features*.

3.4.4 ISO 19115 Geographic Information - Metadata

Der ISO-Standard 19115 *Geographic Information – Metadata* stellt ein Schema zur Beschreibung geografischer Daten und eingeschränkt auch Diensten zur Verfügung. In die Entwicklung von ISO 19115 flossen zahlreiche schon vorhandene internationale Standards mit ein, unter anderem der europäische Prä-Standard ENV 12657 *Geographic Information – Data description – Metadata*, der ANZLIC⁵ Metadatenstandard und vor allem der *Content Standard for Digital Geospatial Metadata (CSDGM)* des FGDC (DANKO 1997).

Inzwischen hat ISO 19115 die drei benötigten Entwurfsphasen durchlaufen und ist als *Draft International Standard (DIS)* vorhanden. Das Dokument beschreibt mehr als 400 Metadatenelemente, die zur Beschreibung von Geodaten herangezogen werden. Unterteilt werden die Elemente in 15 Pakete, die in Abb. 5. dargestellt werden.

Diese Paketstruktur dient der inhaltlichen Unterteilung der Metadatenelemente und darf nicht als Schemadefinition im Sinne eines Datenbankentwurfes verstanden werden. Die Elemente von ISO 19115 werden entweder als verpflichtend (*mandatory*), bedingt (*conditional*) oder wahlweise (*optional*) bezeichnet. Verpflichtende Elemente müssen unbedingt angegeben werden, bedingte Elemente müssen innerhalb eines Schemas auftauchen, falls ein bestimmtes anderes nicht-verpflichtendes Element verwandt wurde und als wahlweise ausgezeichnete Attribute können benutzt werden, müssen aber nicht, um ISO-Konformität zu erreichen. Die *Core metadata elements* stellen 22 Attribute dar, die in jedem Metadatenkatalog der ISO-Konformität erreichen will, implementiert sein müssen. Allerdings sind nicht alle Kernelemente als verpflichtend gekennzeichnet, wichtig ist in Hinsicht auf die freiwilligen Elemente nur, dass der Katalog die Möglichkeit aufweist, das Metadatenelement anzugeben.

⁵ *Australia New Zealand Land Information Council*

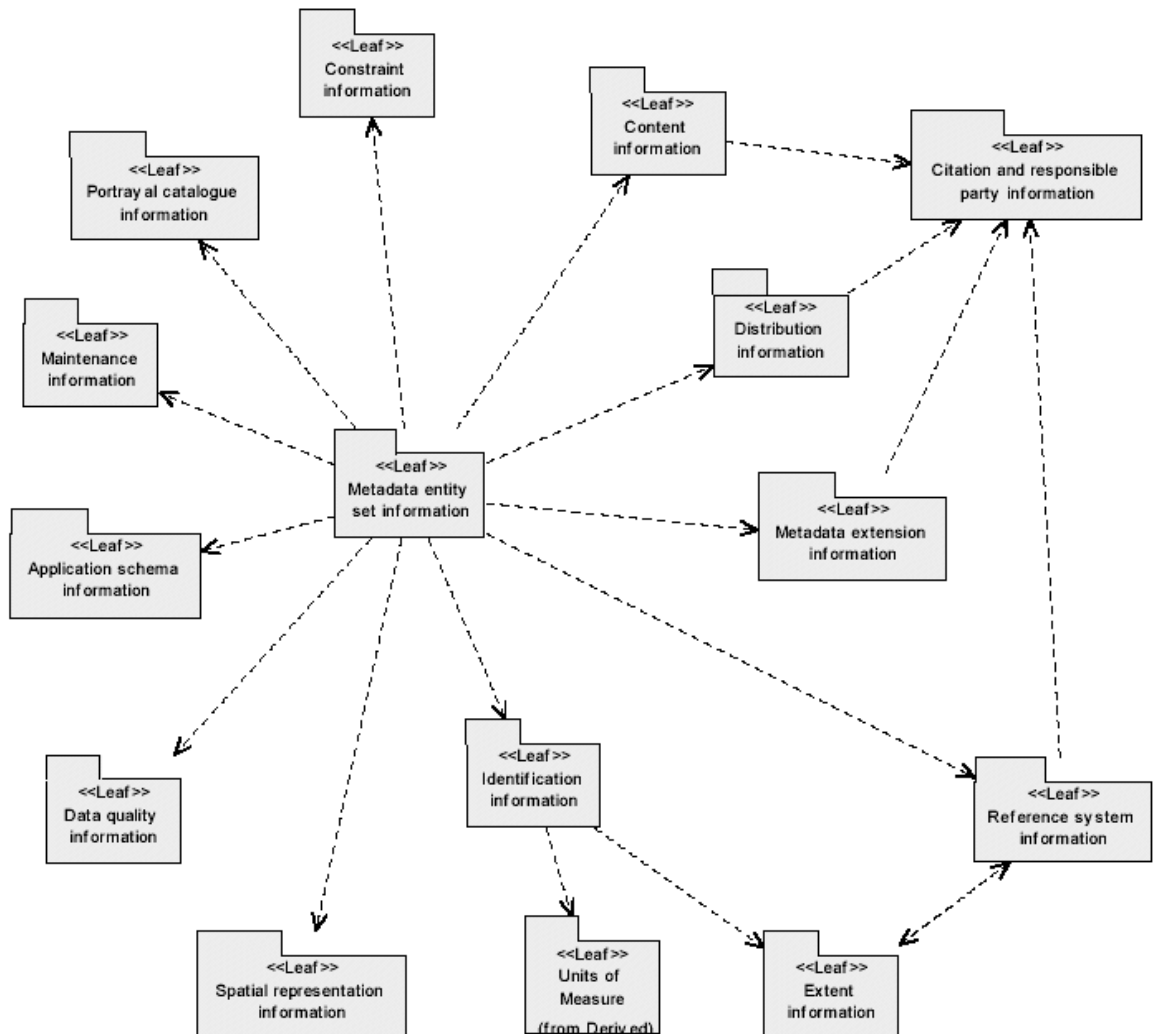


Abbildung 5. ISO 19115 Paketstruktur (Quelle: ISO 2000)

Die Erweiterung des ISO-Schemas durch anwendungsrelevante Elemente wird mittels *Extensions* erlaubt. Solche Erweiterungen können für den Fall definiert werden, dass für ein bestimmtes Anwendungsgebiet ein Metadatenelement benötigt wird, das nicht mittels der schon vordefinierten Elemente abgedeckt ist. *Extensions* können sich auf ein bereits vorhandenes Element beziehen und dieses genauer als in der Vorgabe definieren, oder ein völlig neues Element einführen. Für bestimmte Anwendergemeinschaften kann es notwendig und sinnvoll sein, ein *Profile* zu definieren (Abb. 6.).

Ein *Profile* besteht aus den ISO-Kernelementen, einer Menge anderer ISO-Elemente und einer zusätzlichen Anzahl von *Extensions*. Motivation für die Entwicklung eines solchen *Profiles* ist, dass es für eine

Anwendergemeinschaft gemeinsame Anforderungen an einen Metadatenstandard gibt, die nicht von ISO 19115 vollständig erfüllt werden. Mittels der Schaffung eines *Profiles* wird erreicht, dass die selben Erweiterungen innerhalb der Anwendergemeinschaft genutzt werden und somit Mehrarbeit vermieden wird. Das InGeo-Metadatenformat des deutschen InGeo-Forums (InGeoForum 2000) stellt ein Beispiel eines solchen Profils dar. Ein *profile* muss die 22 Pflichtelemente umfassen und beliebig viele optionale Elemente und *Extensions*.

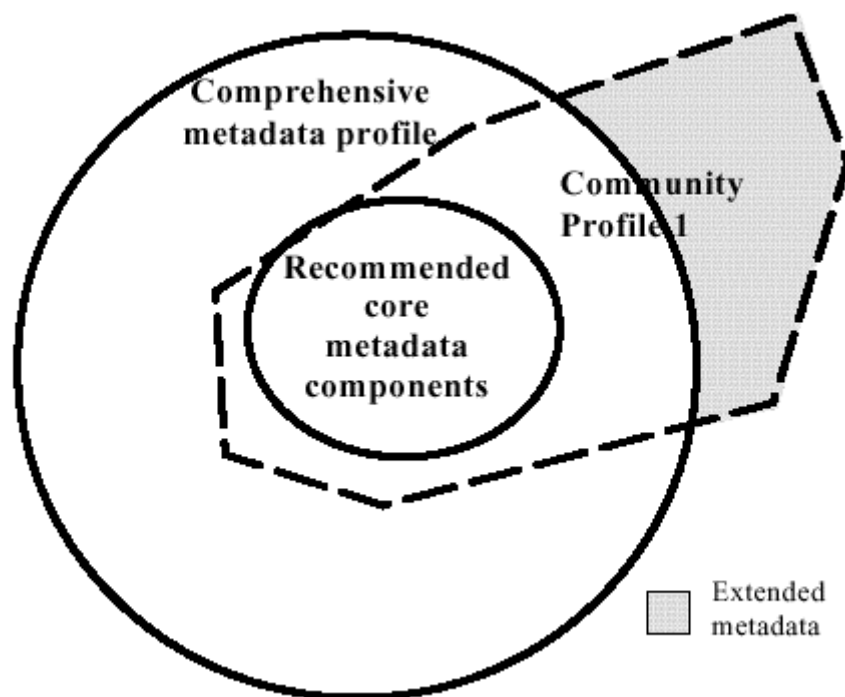


Abbildung 6. ISO 19115 Profile (Quelle: ISO 2000)

ISO 19115 stellt den bislang umfassendsten und flexibelsten Ansatz zur Standardisierung von Geo-Metadaten dar. Die große Menge der zur Verfügung gestellten Elemente deckt mit Sicherheit die Bedürfnisse der meisten Metainformationsverarbeiter ab. Durch die geringe Anzahl an Kernelementen und die Möglichkeit, mittels *Extensions* und *Profiles* einen eigenen, angepassten ISO 19115 Katalog zu erstellen, ist der Standard trotzdem sehr flexibel in seinen Anwendungsmöglichkeiten. Das größte

Hindernis seiner Anwendung stellt allerdings die Komplexität des Dokumentes dar, die Folge seines Umfangs und seiner Flexibilität ist.

3.4.5 Dienstdefinition bei OGC und ISO

Das Dokument ISO 19118 *Geographic Information – Services* (ISO 2001) wurde mit dem *Topic 12* der *Abstract Specification* mit dem Namen *Service Architecture* (OGC 2001a) harmonisiert und hat dieses Dokument inzwischen vollständig ersetzt. Es stellt ein abstraktes Modell für raumbezogene Dienste zur Verfügung. Ein Dienst (*Service*) wird hier verstanden als ein unterscheidbarer Teil der Funktionalität, welcher durch eine Entität mittels Schnittstellen zur Verfügung gestellt wird, wobei eine Schnittstelle eine Menge von Operationen darstellt, die das Verhalten der Entität definieren. Eine Operation wiederum ist die Spezifikation einer Transformation oder Anfrage, die an ein Objekt gesendet werden kann. Dienste können genau wie Datensätze in Katalogen registriert werden, wobei zwischen Diensten unterschieden werden kann, die fest an einen Datensatz gekoppelt sind und solchen, die es nicht sind. Dienste der ersten Art müssen Metadaten über den angekoppelten Datensatz zusätzlich zu den eigentlichen Dienste-Metadaten zur Verfügung stellen, nicht an Datensätze gekoppelte Dienste müssen nur diese Dienste-Metadaten besitzen. Um die Recherche nach katalogisierten Diensten zu ermöglichen, benötigen diese also Metadaten. Aus diesem Grund wird im Rahmen der *OpenGIS Service Architecture* ein Schema für Dienste-Metadaten zur Verfügung gestellt. Die in diesem Schema verwendeten Metadatenelemente sind, soweit möglich, aus ISO 19115 entnommen. Zusätzlich wurden einige Elemente eingeführt, die dienstespezifisch sind, wie beispielsweise der Typ des Dienstes oder die Plattform (DCP⁶), auf welcher der Dienst implementiert wurde.

3.4.6 OpenGIS Katalogdienste

Das *Topic 13* der *Abstract Specification* (OGC 1999d) und die *Catalog Interface Implementation Specification* (OGC 1999e) befassen sich mit der Spezifikation von Geodatenkatalogen. Hierbei werden sowohl Dienste zur

⁶ DCP= *Distributed Computing Platform*

Recherche nach räumlichen Ressourcen beschrieben, als auch Dienste, die den direkten Zugang zu den räumlichen Ressourcen zur Verfügung stellen. Mit räumlichen Ressourcen sind hierbei raumbezogene Dienste oder Geodaten gemeint. Ein Katalog wird hier definiert als ein spezialisiertes Informationssystem für räumliche Ressourcen, die für eine Nutzergruppe zugänglich sind. Die grundlegenden Funktionen, die ein Katalog erfüllen soll, sind:

- Unterstützung bei Organisation und Management verschiedener raumbezogener Daten und Dienste für Recherche und Zugriff,
- Auffinden von Ressourceninformationen aus diversen Quellen und Zusammenführen dieser Informationen an einem Ort sowie
- Möglichkeiten, um diese Ressourcen ausfindig zu machen, auf sie zuzugreifen und sie zu speichern oder, im Falle von Diensten, sie aufzurufen.

Der Begriff Geodatenkatalog oder Geodienstekatalog⁷ ist somit eine Präzisierung des Begriffes Metainformationssystem. Die Architektur der Katalogdienste wird mittels eines Referenzmodells in der *Catalog Interface Implementation Specification* (OGC 1999e) beschrieben (Abb. 7.).

Die hier beschriebene *Application* kommuniziert mit dem *Application Server* mittels der spezifizierten *Catalog Service* Schnittstellen. Dieser Server greift wiederum auf einen Metadatenpeicher (*Metadata Store*) und gegebenenfalls auf einen Datenspeicher (*Data Store*) zurück, um raumbezogene Daten oder Dienste zur Verfügung zu stellen. Wie Abb. 7. zu entnehmen ist, beschreibt der *Metadata Store* den *Data Store*, existiert aber unabhängig davon. Die in der Abbildung erwähnten *Features*-Schnittstellen beschreiben den Zugriff auf die eigentlichen Geodaten im Sinne des OGC. Im Falle eines Kataloges, der nicht nur Geodaten, sondern auch Geodienste katalogisiert, muss das Architekturmodell allerdings erweitert werden. Der *Application Server* würde hierbei an den Nutzer Verweise auf Geodienste als Ergebnis einer Recherche liefern. Im Diagramm müssten demnach zusätzliche Server aufgeführt werden,

⁷ innerhalb des OGC wird manchmal auch von *Web Registry Services* gesprochen

welche die katalogisierten Dienste zur Verfügung stellen, die dann wiederum auf verschiedene Datenquellen zugreifen.

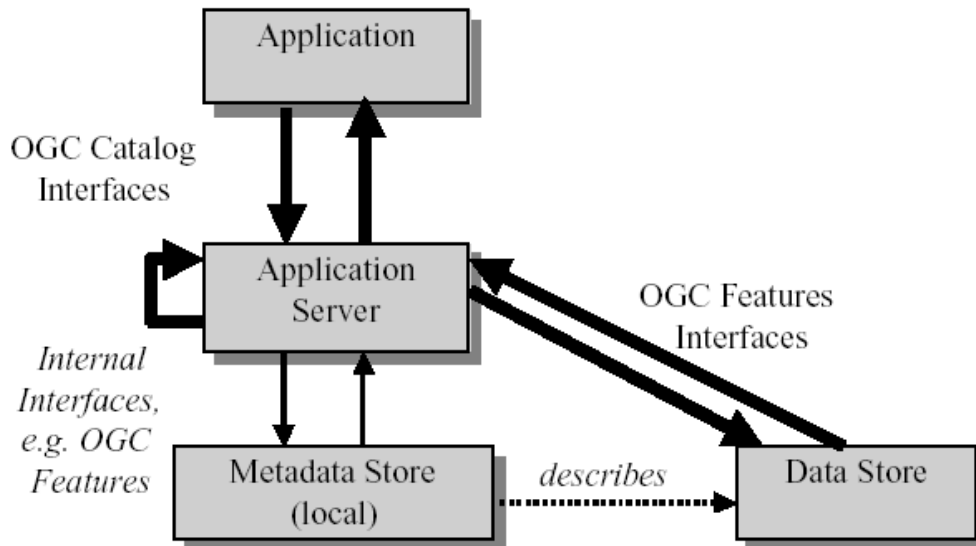


Abbildung 7. Architektur des Referenzmodells (Quelle: OGC 1999e)

Die *Catalog Interface Implementation Specification* beschreibt in der Version 1.0 neben dem abstrakten Modell Schnittstellen für drei Plattformen, welche in Form von Profilen zur Verfügung gestellt werden. Diese sind die *Common Object Request Broking Architecture* (CORBA), das *Common Object Model* (COM) und Z39.50, ein Protokoll zur Definition verteilter Datenbankanfragen. Die Unterteilung der verschiedenen Dienste, die von einem Katalog zur Verfügung gestellt werden, ist in Abb. 8. dargestellt. Hiernach kann ein Katalogdienst entweder dem Auffinden, dem Zugriff oder dem Management von Informationsressourcen dienen. *Discovery Services* sind Dienste, die einem Client erlauben, Daten oder Dienste beschreibende Metadaten zu benutzen, um Informationsressourcen zu entdecken. Dies geschieht mittels einer Anfragesprache, die beispielsweise ein SQL-Dialekt sein kann, der räumliche Operatoren unterstützt oder die *Common Catalog Query Language* (CQL), welche Teil der *Catalog Service Interface Definition* ist. *Access*

Services erlauben den Zugriff auf Geodaten oder Dienste, entweder direkt oder über einen außerhalb des Katalogdienstes ablaufenden Weg⁸.

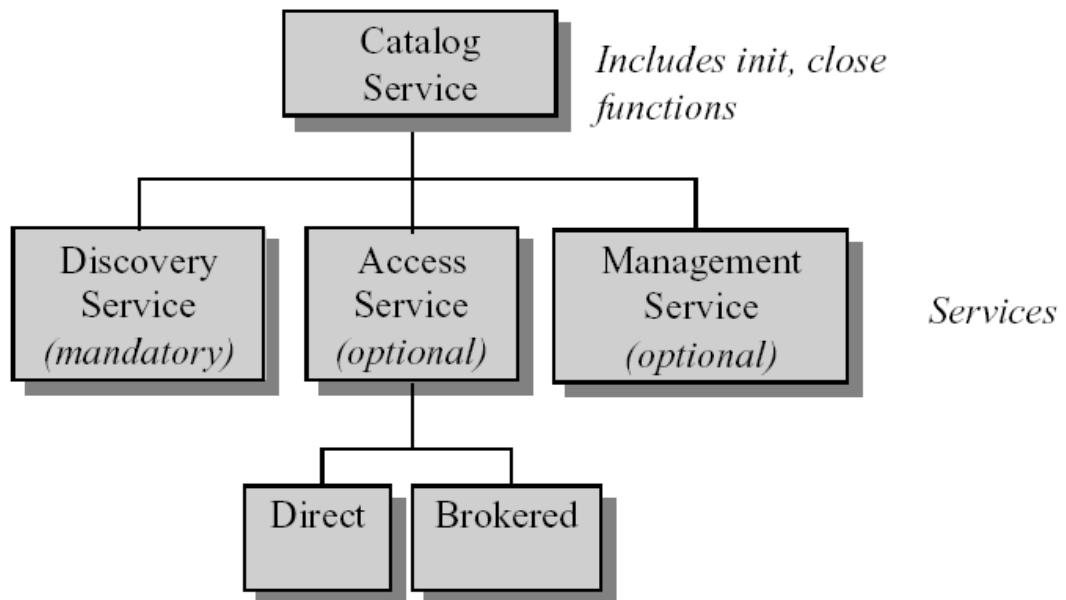


Abbildung 8. Zerlegung der Katalogdienste (Quelle: OGC 1999e)

Die *Catalog Interface Implementation Specification* ist unabhängig von einem bestimmten Metadatenchema. Es wird jedoch darauf verwiesen, dass ein OGC-konformer Geodatenkatalog die Kernelemente von ISO 19115 benutzen sollte. Die Verwendung des Begriffes Metadaten im Sinne der Spezifikation umfasst Vorschaubilder⁹ und andere Formen von nicht-textueller Art von Metadaten, bis hin zu dynamisch aus den Originaldaten abgeleiteten Vorab-Karteninformationen.

Das in der *Catalog Interface Implementation Specification* entworfene Objektmodell ist Abb. 9. zu entnehmen. Dieses Diagramm ist deshalb besonders interessant, weil das gesamte konzeptionelle Metadatenmodell des OGC darin abzulesen ist.

Ein OpenGIS *Catalog Service* benutzt einen *Catalog*, um auf Metadaten und Daten zuzugreifen und an Anfragende Auskunft zu geben. Dieser *Catalog* besteht aus *Catalog Entries* und diese wiederum aus *MetadataEntities*. Ein

⁸ was dem *Brokered Access Service* entspricht

⁹ es wird hierbei von *Browse Images* gesprochen

CatalogEntry ist einem Datensatz zugeordnet und beschreibt diesen. Interessant an diesem Modell ist, dass die Klasse *CG_Catalog* eine Spezialisierung der Klasse beziehungsweise der Schnittstelle *FeatureCollection* darstellt und *CG_CatalogEntry* von *Feature* erbt beziehungsweise das *Feature Interface* implementiert. Dies bringt zum Ausdruck, dass ein Eintrag in einem Metadatenkatalog ein Objekt mit Raumbezug ist und somit im Sinne des OGC wiederum ein *Feature* darstellt. Ein einzelnes Metadatenelement wird somit zum Attribut eines *Features*, ein Konstrukt, das schon im Zusammenhang mit dem *Essential Model* vorgestellt wurde.

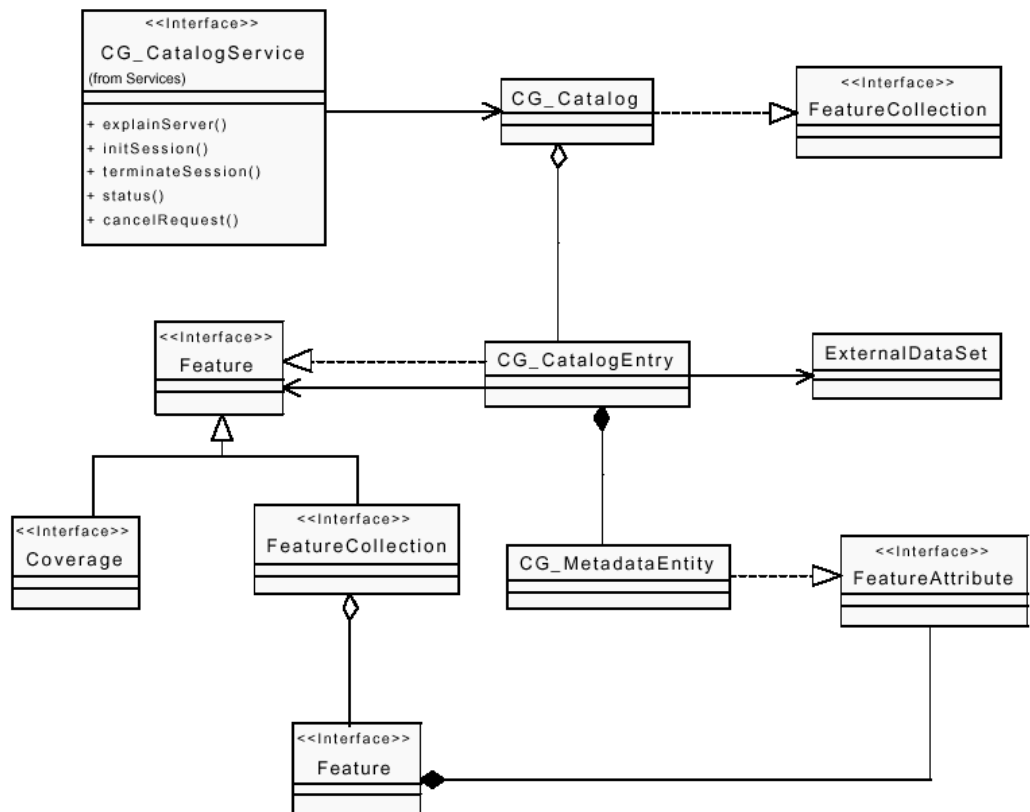


Abbildung 9. Katalog Objektmodell (Quelle: OGC 1999e)

Da Katalogdienste nicht nur zur Recherche nach Geodaten, sondern auch zur Katalogisierung von raumbezogenen Diensten dienen sollen, ist die *Catalog Interface Implementation Specification* dieses Thema betreffend gerade in Überarbeitung. Ein sich bislang noch im Entwurfsstadium befindliches Dokument, das den Namen *Web Services Stateless Catalog Profile* (REICH &

PANAGIOTIS 2001) trägt, stellt in erster Hinsicht eine Erweiterung der *Catalog Interface Implementation Specification* um eine weitere Plattform dar. Diese Plattform ist das *World Wide Web* und das benutzte Protokoll das *Hypertext Transfer Protocol* (HTTP). Viel interessanter ist dieses Dokument im Hinblick darauf, dass hier konsequent nicht nur die Katalogisierung von Geodaten, sondern auch die von Geodiensten beschrieben wird. Die hiermit implementierten Kataloge, die manchmal auch als *Web Registries* bezeichnet werden, beinhalten sowohl Metadaten zur Beschreibung von Diensten als auch von Datensätzen. Bezüglich Dienste beschreibender Metadaten wird in diesem Dokument auf die entsprechenden Teile von ISO 19119 verwiesen. Das *Web Services Stateless Catalog Profile* wird in die Version 2 der *Catalog Interface Implementation Specification* integriert werden.

3.5 Zusammenhang des *OpenGIS* Modells und der semiotischen Theorie

Das *OpenGIS* Modell von Metadaten auf Ebene des *Essential Models* und des *Abstract Models* unterscheidet Metadatensätze und Metadatenelemente, die entweder mit einem oder mehreren *Features* oder *FeatureCollections* assoziiert sein können. Eine darüber hinausgehende Unterscheidung wird nicht vorgenommen, beispielsweise in Hinblick auf verschiedene Arten von Metadaten, weder in den *Abstract Specifications*, noch in den Katalogdiensten spezifizierenden Dokumenten. Ebenso wird in ISO 19115 keine diesbezügliche Unterscheidung vorgenommen, sondern alle möglichen Metadatenelemente aufgelistet. Die in ISO 19115 vorgenommene Beschreibung von Metadatenelementen ist unabhängig von der Nutzung der Elemente. Es wird in der Einleitung ausdrücklich darauf hingewiesen, dass die vorgeschlagenen Elemente für die verschiedensten Zwecke benutzt werden können (ISO 2000).

Das in dieser Arbeit entwickelte, auf den Grundlagen der Semiotik aufbauende Modell stellt somit eine Erweiterung des OGC Modells dar. Da innerhalb dieser Arbeit bereits die Terminologie des OGC benutzt wurde, existiert kein Widerspruch zwischen den beiden Modellen. Das semiotische Modell kann als

Kommentar und Verfeinerung zu den Spezifikationen des OGC verstanden werden.

3.6 Vorgehensweise

In den vorangegangenen Abschnitten von Kapitel 3 wurde der Objektbereich und Referenzrahmen dieser Untersuchung abgesteckt. Innerhalb dieses Referenzrahmens wird durch Verknüpfung und Anwendung semiotischer Konzepte auf Phänomene raumbezogener Informationsverarbeitung ein allgemeines Modell raumbezogener Metainformationsverarbeitung erzeugt. Um diese Modellerzeugung vorzubereiten, werden im nächsten Kapitel aktuelle Ansätze der Metainformationsverarbeitung vorgestellt, wie sie in verschiedenen Disziplinen und inhaltlichen Bereichen durchgeführt werden. Diese Ansätze verdeutlichen den Stand der Forschung im Bereich Metainformation und werden unter Berücksichtigung der durch das OGC definierten Setzungen interpretiert und analysiert. Diese Analyse bildet die Grundlage der darauf folgenden Ausführungen.

Nach diesem allgemeinen und somit nicht geo-spezifischen Überblick beginnt die Entwicklung des allgemeinen Modells der raumbezogenen Metainformationsverarbeitung. Vorangestellt werden hierfür einige Setzungen für Definitionen, welche die Grundbegriffe der weiteren Analyse darstellen, wie beispielsweise Information oder Zeichen. Diese Definitionen sind aus der Literatur entnommen beziehungsweise abgeleitet worden, stellen aber keine selbständige Weiterentwicklung der gängigen Konzepte dar. Auf diesen Setzungen aufbauend werden dann Phänomene raumbezogener Metainformationsverarbeitung aus Sicht der Semiotik interpretiert, um so zu Einsichten über ihre Funktion zu gelangen. Um die geistige Brücke zurück zur *GIScience* zu schlagen, werden die semiotisch motivierten Funktionen mit Begriffen der Geoinformationsverarbeitung belegt. Ergebnis dieses Vorgehens ist ein tieferes Verständnis und eine fundierte Terminologie der Geo-Metainformationsverarbeitung.

4 Aktuelle Ansätze zur Metainformationsverarbeitung

Die Aufgabe dieses Kapitels ist es, den derzeitigen Stand von Forschung und Praxis der Metainformationsverarbeitung unter besonderer Berücksichtigung der raumbezogenen Problematik darzustellen. Zu Beginn wird der Stand der Forschung diskutiert, beziehungsweise anhand von Referenzen zu verschiedenen Teilen der Arbeit, die sich mit dem aktuellen Forschungsstand auseinandersetzen, dokumentiert. Danach werden Ansätze zur Metainformationsverarbeitung in verschiedenen Disziplinen vorgestellt, um das Verständnis des Problembereichs zu vertiefen und die Aufgaben, die Metainformationen zu lösen im Stande sind zu identifizieren. Ein Vergleich der verschiedenen Ansätze und eine Klassifikation der Funktionen von Metainformation bildet den Abschluss des Kapitels.

4.1 Stand der Forschung

Die Literatur zur speziell raumbezogenen Metainformationsverarbeitung ist relativ überschaubar, zudem sind viele der theoretischeren Aufsätze zu diesem Thema im Kontext der Umweltinformationsverarbeitung entstanden. Unterschiedliche Aspekte des Standes der Forschung werden in verschiedenen Teilkapiteln dieser Arbeit in ihrem jeweiligen Kontext diskutiert und sollen hier nur referenziert werden. Abschluss dieses Abschnittes bildet ein Überblick über die Forschungsbereiche Ontologien und Wissensmanagement, die beide zum Gegenstand der vorliegenden Arbeit benachbart sind und komplementäre Aufgaben erfüllen. Die Bedeutung von Ontologien in Beziehung zur hier behandelten Aufgabe wird in diesem Zusammenhang gleichfalls erörtert.

In Kapitel 3 dieser Arbeit werden die praxisrelevanten Aspekte der Geo-Metainformationsverarbeitung im Umfeld des *Open GIS Consortiums* diskutiert. Der hauptsächliche Fokus bei dieser Diskussion spielt die im OGC eine zentrale Stellung einnehmende Interoperabilitätsfrage, die mittels Standardisierung gelöst werden soll.

In den Abschnitten 4.2 bis 4.10 werden die derzeitigen praktischen Ansätze der Metainformationsverarbeitung diskutiert. Dies schließt weite Teile des Forschungsstandes mit ein. Abschnitt 4.5 im Speziellen gibt einen Überblick über die aktuellen Initiativen der Umwelt-Metainformationsverarbeitung. Vor allem der UDK stellt eines der fortgeschrittensten und theoretisch fundiertesten Projekte zur Katalogisierung von Umweltdaten dar und spiegelt somit auch den Forschungsstand wider. In Abschnitt 4.6 werden die wichtigsten Initiativen der Geo-Metainformationsverarbeitung vorgestellt. Hier stellen vor allem die Initiativen GDI NRW und GeoMIS.Bund den Stand der aktuellen angewandten Forschung dar.

In Kapitel 5 wird ein auf der Semiotik aufbauendes Modell der raumbezogenen Metainformationsverarbeitung entwickelt. Hauptsächlich in Abschnitt 5.8, in dem die wichtigsten Definitionen entwickelt werden, findet eine Auseinandersetzung mit anderen Autoren, die sich mit diesem Thema beschäftigt haben, statt. Am Ende des Abschnittes 5.8 werden verschiedene andere Klassifikationen mit dem in dieser Arbeit entwickelten Ansatz verglichen.

Ein aktuelles und mit dem Forschungsgegenstand Metainformation eng verwandtes Forschungsfeld stellen Ontologien dar. Ontologien sind „[...]Theorien, die ein spezifisches Vokabular benutzen, um Entitäten, Klassen, Eigenschaften und Funktionen zu beschreiben, die in Beziehung zu einer speziellen Sicht der Welt stehen“ (übersetzt aus FONSECA 2000, S. 1). Sie stellen somit die Formalisierung von semantischen Informationen dar. Mittels der Entwicklung und Anwendung von Ontologien verspricht man sich die Lösung semantischer Heterogenitätsprobleme (HAKIMPOUR & TIMPF 2001).

Wie bereits erwähnt, formalisieren Ontologien die Semantik eines bestimmten Ausschnittes aus der Welt, der für eine bestimmte Gruppe von Personen von Interesse ist. Gemeinsam entwickelte Ontologien erlauben es, Personengruppen semantische Übereinstimmung bei der Beschreibung ihrer Daten zu erreichen und ermöglichen so semantische Interoperabilität.

Wenngleich Überschneidungen zwischen den Aufgabenbereichen von Metadaten und Ontologien bestehen, sind die beiden Ansätze als komplementär zu betrachten. Ontologien lösen semantische

Interoperabilitätsprobleme und erlauben es, beispielsweise in Form eines Thesaurus, Verbindungen zwischen den semantischen Welten verschiedener Nutzergruppen von Geoinformationen zu erstellen. Mittels Ontologien alleine kann man jedoch keine Daten auffinden oder über ihre allgemeine Einsetzbarkeit für eigene Zwecke entscheiden. Solche Aufgaben werden von Geodatenkatalogen übernommen, die hierfür Metainformationen benötigen. Ontologien können allerdings zur Unterstützung des Recherchevorgangs in Katalogen benutzt werden, indem sie die semantische Begrifflichkeit des Suchenden mit der Semantik des Entstehungskontextes der Daten verbinden. Aus diesen Gründen und aus dem einfachen Grund, dass Ontologien bislang in der Praxis der raumbezogenen Metainformationsverarbeitung noch keine nennenswerte Rolle spielen, werden Ontologien innerhalb dieser Arbeit keine weitere Beachtung finden.

Ein weiteres aktuelles Forschungsfeld, das eng mit den in dieser Arbeit behandelten Problemen verwandt ist, ist das Wissensmanagement¹⁰. Wissensmanagement befasst sich mit der organisatorischen Seite von Daten- und Informationsverarbeitung und dem Zusammenspiel zwischen den Technologien, die diese Aufgaben bewältigen, und dem Menschen (MALHOTRA 1998). Der Fokus der hier getätigten Überlegungen liegt vor allem darin, wie menschliche Kreativität und Informationstechnologie im Zusammenspiel zu möglichst effizienten Problemlösungen gelangen. Unter dem Dach des Wissensmanagements befasst man sich mit Ansätzen wie dem *Data Warehousing* (siehe auch Abschnitt 4.4), der Verwaltung von Metainformationen, dem Einsatz von Ontologien (IUK 2000) und dem *Content Management* (ARNDT, GÜNTHER & PLUSKAT 2001), aber vor allem damit, wie man mit Hilfe dieser Ansätze Geschäftsprozesse optimieren kann und welchen wirtschaftlichen Nutzen diese Technologien hervorbringen (MALHOTRA 2002). Die vorliegende Arbeit befasst sich mit Theoriebildung für raumbezogene Metainformation, um das Management von Geodaten mittels Geodatenkatalogen und Geodienstekatalogen zu verbessern und kann somit als Beitrag zum *Knowledge Management* gesehen werden.

¹⁰ im englischen *Knowledge Management*

4.2 Ansätze zur Metainformationsverarbeitung in verschiedenen Disziplinen

Metainformationsverarbeitung wird in verschiedenen Problemkontexten betrieben. Klassischerweise wird im Datenbank-Umfeld mit Metadaten gearbeitet und somit auch in verwandten Gebieten wie dem *Data Warehousing*. Daneben sind Ansätze im Bibliothekswesen und im Internet von zentraler und grundlegender Bedeutung und finden deshalb in dieser Arbeit besondere Beachtung. Speziell das Bibliothekswesen besitzt eine lange Tradition um Umgang mit Metainformationen. Ansätze aus der Umwelt- und Geoinformationsverarbeitung werden hier diskutiert, da das Ziel der Arbeit kein allgemeines Modell der Metainformationsverarbeitung ist, sondern raumbezogene Metadaten im Mittelpunkt der Arbeit stehen. Außer diesen Bereichen gibt es Ansätze zur Metainformationsverarbeitung in einer Vielzahl inhaltlicher Domänen, wie der Mathematik, der Physik, der Chemie und der Medizin. Die Medizininformatik wurde hier als ein exemplarisches Beispiel für diese Vielzahl von Domänenansätzen gewählt. Um zu einer Annäherung an den Begriff der Metainformation zu gelangen, sollen hier die wichtigsten Ansätze aus den genannten Bereichen diskutiert werden.

Es ist anzumerken, dass die folgende Kategorisierung keine überlappungsfreien Mengen von Anwendungsfällen hervorbringt, sondern große Überschneidungsbereiche bestehen. In besonderem Maße sollen international weit verbreitete oder in Deutschland wichtige Ansätze berücksichtigt werden. Die folgende Aufzählung erhebt in keiner Weise den Anspruch auf Vollständigkeit, da die Zahl der Metainformationssysteme und Standards einen solchen Anspruch inzwischen nahezu unmöglich macht. Es soll aber mittels dieser Vorstellung gezeigt werden, auf welche Weise mit Metainformationen umgegangen wird und welche grundlegenden Funktionen sie übernimmt.

Viele der vorgestellten Ansätze benutzen inzwischen die Datenauszeichnungssprache *eXtensible Markup Language* (XML) zur Datendefinition und zum Datenaustausch. Die Entwicklung von XML ist eng mit dem Internet verknüpft, hat sich aber inzwischen in allen Bereichen der informationsverarbeitenden Technologie durchgesetzt. Wegen dieser

domänen- und disziplinübergreifenden Nutzung von XML spielt diese Technologie eine besondere Rolle für die Metainformationsverarbeitung. Die grundlegenden Funktionalitäten und Eigenschaften von XML werden im Abschnitt über das Internet (Abschnitt 4.8) erklärt.

4.3 Datenbanken

Im Datenbank-Umfeld werden Metadaten benutzt, um Informationssysteme und Datenbanken zu verwalten. Der Bereich der Datenbank, der Metadaten enthält, wird manchmal auch als *Data Dictionary*, Datenkatalog oder *Metadaten-Repository* bezeichnet (HEUER & SAAKE 1997). Hier werden technische Metadaten, die zur Verwaltung der Datenbank notwendig sind, vorgehalten, das heißt Informationen, die das Schema einer Datenbank beschreiben, wie beispielsweise den Aufbau von Relationen und Datentypen. Darüber hinaus werden in einem *Data Dictionary* Definitionen von Sichten, gespeicherte Prozeduren und Statistiken zur Datenbankoptimierung verwaltet. Im Gegensatz zu Integritätsbedingungen definieren Metadaten in einem *Data Dictionary* einzelne erlaubte Datenbankzustände und nicht die Gesamtmenge von erlaubten Zuständen (BELL 1997). Ein Datenkatalog einer Datenbank besitzt zwei Aufgaben. Er dient in erster Hinsicht dem Datenbankmanagementsystem (DBMS) selbst zur Speicherung der Daten, die benötigt werden, um die Datenbank zu verwalten. Metadaten in *Datenbank-Repositories* sind somit primär nicht zur Nutzung durch den Menschen vorgesehen, sondern enthalten automatisiert verarbeitbare Daten. Dies bedeutet allerdings nicht, dass diese Metadaten für Menschen nutzlos sind. Sie können ebenfalls benutzt werden, um tatsächlich gespeicherte Daten zu identifizieren, oder sie können zur Diagnose bei Leistungsproblemen des DBMS herangezogen werden. Diese Metadaten werden wie alle anderen Daten im Datenmodell des DBMS realisiert. Um den Zugang zu diesen verteilten Metadaten benutzerfreundlicher zu gestalten, gibt es die Möglichkeit, zusätzliche Sichten auf die Metadaten zu definieren, eine Vorgehensweise, die von den meisten DBMS unterstützt wird (HEUER & SAAKE 1997).

Ein Datenkatalog, wie er im Datenbankkontext bezeichnet wird, hat somit gewisse Unterschiede zu einem Katalog im OGC-Jargon, über die Differenzierung zwischen räumlichen und nicht-räumlichen gespeicherten Daten hinausgehend. Dieser Unterschied liegt im Zweck, dem ein solcher Katalog dient. Ein OGC-Katalog dient dem Zweck, die Suche nach Informationsressourcen zu unterstützen. Ein solcher Katalog kann zwar trotzdem gänzlich technische Metadaten verwalten, sie werden aber immer zur Recherche oder eventuell zum Datenzugriff eingesetzt. Im Gegensatz dazu wird ein Datenbank-Katalog von einem DBMS benötigt, um eine Datenbank zu verwalten. Erst in zweiter Hinsicht dient er Personen zur Recherche nach Informationen.

4.4 Data Warehouse

Data Warehouses sind Datenbanken, die historische Daten integrieren und zusammenfassen und Daten aus verschiedenen Quellen referenzieren (HALLMARK 1995). Der Ansatz zur Metadatenverarbeitung ist somit von der Datenbank-Sichtweise abgeleitet. Da die wesentliche Aufgabe von *Data Warehouses* die Integration von Daten ist, die nicht notwendigerweise innerhalb einer Datenbank gehalten werden, kommt der Metadatenverarbeitung hier eine zentrale Rolle zu. Sie finden ihren Einsatz oftmals in größeren Firmen, die verteilte und heterogene Datenbestände besitzen. *Data Warehouses* sind selbst große *Repositories*, in die verschiedene Datenquellen zum Zwecke der übergreifenden Analyse integriert werden. Analysen können beispielsweise aufbauend auf der Zeit, nach Produkten, die von einer Firma erstellt werden, oder nach räumlicher Verbreitung durchgeführt werden, um Trends zu identifizieren und somit Wettbewerbsvorteile zu sichern (HALLMARK 1995).

Metadaten von guter Qualität sind von großer Bedeutung für ein *Data Warehouse*. Sie werden während der Datenerzeugung, der Datentransformation und der Datenabfrage benutzt (MINNESOTA HISTORICAL SOCIETY 2000). Metadaten, die bei der Datenerzeugung angelegt werden, beinhalten beispielsweise die Herstellungsgeschichte der Daten und die Beschreibung der Algorithmen, die zur Verarbeitung der Daten benutzt

wurden. Während der Datentransformation entstehen Metadaten, die den Ablauf der Datentransformation, Namensänderungen oder Änderungen anderer Merkmale aufführen. Bei der Datenabfrage ist Metainformation notwendig, um die Navigation in den Datenbeständen zu unterstützen.

Außer diesen drei Typen von Metadaten benötigt ein *Data Warehouse* zusätzlich Informationen über den Umfang der Nutzung, das Wachstum des Datenvolumens in Beziehung zum zur Verfügung stehenden Speicherplatz und Informationen über Nutzerrechte, sprich wiederum verschiedene Metainformationen (MINNESOTA HISTORICAL SOCIETY 2000).

Offensichtlich spielen Metadaten eine wichtige Rolle beim *Data Warehousing*. Von zentraler Bedeutung in diesem Bereich ist die Integration von Daten aus verschiedenen Quellen. Dies setzt voraus, dass Entstehungskontext und technisch-strukturelle Informationen gut dokumentiert sind. Katalogisieren und Auffinden von nicht bekannten Datenbeständen sind hier nur von geringerer Bedeutung im Gegensatz zur Integration und Analyse von Daten verschiedener Quellen.

4.5 Umweltinformationssysteme

Ein Umweltinformationssystem (UIS) gibt Auskunft über den Zustand von Gewässern, Luft, Boden sowie biologischer Spezies und ihrer Lebensräume. Darüber hinaus werden auch Tätigkeiten, die den Zustand der Umwelt beeinträchtigen oder schützen, als Umweltinformationen bezeichnet (GREVE 1999) und somit in Umweltinformationssystemen verarbeitet¹¹. Moderne Umweltinformationssysteme sind keine geschlossenen Systeme, sondern werden aus Komponenten aufgebaut. Eine notwendige Komponente jedes UIS ist ein Metainformationssystem. Eines der wichtigsten Probleme bei der Verarbeitung von Umweltinformationen ist deren Integration und Verfügbarmachung, da sich Umweltinformationen durch große Heterogenität und oftmals sehr dezentrale Speicherung auszeichnen. Viele öffentlich zugängliche Umweltinformationssysteme sind reine

¹¹ eine detailliertere, auf dem deutschen Umweltinformationsgesetz aufbauende Definition erfolgt in Abschnitt 5.7

Metainformationssysteme, die sich somit nur mit der Auskunft über Informationsquellen befassen.

Der Aufbau von Umweltinformationssystemen wird inzwischen verstärkt durch zahlreiche Gesetzgebungen, wie z.B. dem deutschen Umweltinformationsgesetz oder der *Council Directive on the freedom of access to information on the environment* der Europäischen Gemeinschaft vorangetrieben. Die verschiedenen Ansätze mündeten in der Aarhus-Konvention, die im Dezember 1998 auch von der Bundesregierung gezeichnet, allerdings noch nicht in deutsches Recht integriert wurde. Diese Gesetzgebungen sehen neben vielen anderen Aspekten vor, dass Umweltinformationen der öffentlichen Verwaltungen für Bürger frei zugänglich gemacht werden müssen. Das Internet als frei zugängliches Kommunikationsmedium ist hierbei als Hauptplattform in Erscheinung getreten. Aus diesem Grund sind alle hier vorgestellten Ansätze der Umwelt-Metadatenverarbeitung mit einem Zugriff über Internet ausgestattet. Innerhalb von Umweltinformationssystemen werden Metainformationssysteme als Integrationskomponenten diverser Fachinformationssysteme benutzt (AUGSTEIN & GREVE 1995).

In Deutschland existieren Umweltinformationssysteme auf verschiedenen Verwaltungsebenen und darüber hinaus auch in Firmen als betriebliche Umweltinformationssysteme (BUIS). Es gibt zahllose Systeme zur Umweltinformationsverarbeitung im kommunalen Einsatz, auf Länderebene und in privatwirtschaftlichen Institutionen. Von übergeordnetem Interesse sind der Umweltdatenkatalog und das Umweltinformationsnetz Deutschland (GEIN).

4.5.1 Umweltdatenkatalog

Das Referenzprojekt für umweltbezogene Metainformationssysteme ist der deutschsprachige Umweltdatenkatalog (UDK) zu dem international gesehen, vor allem im Hinblick auf seinen flächendeckenden Einsatz, kein vergleichbares Gegenstück existiert. Der UDK ist ein Metainformationssystem für umweltrelevante Daten in öffentlichen Verwaltungen und wurde von 1991 bis 1995 im niedersächsischen Umweltministerium, auf der Basis einer

Länderkooperation und eines Forschungs- und Entwicklungsvorhabens des Umweltbundesamtes (UBA), entwickelt (GÜNTHER ET AL. 1996). Der UDK wird von den Umweltverwaltungen Mehrzahl der deutschen Bundesländer¹², einigen Landesvermessungsverwaltungen und verschiedenen Bundesinstitutionen eingesetzt. In der Republik Österreich ist der Einsatz des Umweltdatenkataloges seit dem 1.1.1995 im Umweltinformationsgesetz vorgeschrieben.

Obwohl das Datenmodell des UDK mit 115 Elementen recht umfangreich ist (NIEDERSÄCHSISCHES UMWELTMINISTERIUM 1998), werden bestimmte Bereiche, die für eine präzise Beschreibung von Geodaten notwendig wären, von ihm nur unvollständig abgedeckt. Vor allem die Ungenauigkeit der im UDK gespeicherten Informationen und die Inkonsistenz in der Befüllung zwischen den verschiedenen ihn einsetzenden Behörden wirft einige Probleme auf. Die Ursache diesen Problems liegt vor allem darin, dass der UDK verhältnismäßig viele Freitextfelder umfasst, die weiten Interpretationsspielraum erlauben. Der UDK, beziehungsweise seine auf Geoinformationen zutreffenden Datenfelder, sollen in naher Zukunft an den ISO-Standard 19115 *Geographic Information – Metadata* (OGC 2000) angepasst werden, was mit hoher Wahrscheinlichkeit zumindest zu einer Teillösung dieser Probleme führen und zu seiner weiteren Verbreitung beitragen wird.

Den UDK gibt es in mehreren technischen Umsetzungen: außer einer Windows-Version existiert im Internet der Virtuelle UDK¹³, der die Recherche in den Datenbeständen des Bundes und denen von sieben Bundesländern erlaubt. Der Zugang zum Virtuellen UDK ist öffentlich (NIKOLAI ET AL. 1999). Der UDK wird fortlaufend weiterentwickelt und an neue Anforderungen angepasst. Die aktuell in Entwicklung befindliche Version 4.3 wird beispielsweise die Nutzung von SOAP-Technologie¹⁴ einführen (Kruse & Eichler 2001).

¹² im Moment alle Bundesländer außer Berlin (siehe http://www.umweltdatenkatalog.de/koudk/kst_partner.html)

¹³ zu finden unter <http://www.umweltdatenkatalog.de>

¹⁴ *Simple Object Access Protocol*, eine Möglichkeit mittels *XML Remote Procedure Calls* durchzuführen

Der UDK ist ein reines Metainformationssystem, das keinen direkten Zugang zu von ihm katalogisierten Originaldaten erlaubt. Der UDK ist unter Benutzung der OGC-Vokabeln ein Datenkatalog, der einen *Discovery Service* zur Verfügung stellt, aber keinen *Access Service*. Die katalogisierten Objekte des UDK sind allerdings nicht nur Geo-Informationsressourcen, sondern umfassen auch andere Entitäten. Die Metainformation des UDK kann auf Entitäten der folgenden Fachbezüge verweisen:

- Organisationseinheit / Fachaufgabe
- Datensammlung / Datenbank
- Dienst / Anwendung / Informationssystem
- Dokument / Bericht / Literatur
- Geo-Information / Karte
- Vorhaben / Projekt / Programm

Für jede dieser Kategorien werden spezifische Datenfelder angelegt. Darüber hinaus existieren allgemeine Metadaten, die zu jedem im UDK verwalteten Objekt angelegt werden. Diese sind in die folgenden Gruppen unterteilt:

- Identifikation
- Verschlagwortung
- Verfügbarkeit
- Zusatzinformationen
- Verweise
- Raumbezug
- Zeitbezug
- Fachbezug

Unter der Überschrift Raumbezug kann für jede beschriebene Entität ein Raumbezug abgelegt werden¹⁵ und zwar durch Angabe eines umgebenden Rechteckes in Gauß-Krüger-Koordinaten, Beschreibung des Rechteckes als fachliche Gebietseinheit und Angabe der zugehörigen administrativen Einheit. Es ist also möglich, für jedes UDK-Objekt einen direkten oder indirekten räumlichen Bezug anzulegen, wodurch das Objekt zu einer Geoinformation wird. Dies trägt wiederum zur Unterstützung der Aussage bei, dass der

¹⁵ die Datenfelder sind optional

Raumbezug in der Umweltinformationsverarbeitung eine große Rolle spielt. Entitäten, die in den Fachbezug Geoinformation/Karte einzuordnen sind, werden darüber hinaus durch Attribute zur Darstellung, Datenqualität und Genauigkeit beschrieben, folglich klassischen Geo-Metadaten.

4.5.2 Umweltinformationsnetz Deutschland

Im Zusammenhang von Metainformation und Umweltinformation in Deutschland ist das Umweltinformationsnetz Deutschland (GEIN¹⁶) unbedingt zu erwähnen, welches seit dem 9. Juni 2000 im Internet verfügbar ist¹⁷ (UBA 2000a). Es ist in einer Bund/Länderkooperation im Rahmen eines Forschungs- und Entwicklungsvorhabens des Umweltforschungsplans (UFOPLAN) beim UBA entstanden.

Bei GEIN handelt es sich nicht um ein Metainformationssystem im engeren Sinne, sondern um eine auf Umweltdaten spezialisierte Suchmaschine. Von GEIN wird dann auf andere Informationssysteme und Informationsressourcen wie die Umweltdatenkataloge der Länder verwiesen. Der UDK ist somit Teil des Umweltinformationsnetzes Deutschland. GEIN referenziert umweltrelevante Daten aus Behörden und öffentlichen Einrichtungen. Ziel von GEIN ist es nicht, neue Informationsquellen zu erzeugen, sondern schon im Internet existierende zu erschließen und recherchierbar zu machen (BANDHOLTZ, BÖS & RÜTHER 2000).

Das Verweissystem GEIN katalogisiert Webseiten, die von den Anbietern von Umweltinformationen zur Verfügung gestellt werden. Eines der zur Recherche herangezogenen Suchkriterien ist das *area*-Attribut, welches den Raumbezug des Dokuments mit einem Deskriptor des Geografischen Thesaurus Umwelt (GTU) beschreibt (UBA 2000b), es wird somit nur indirekte Georeferenzierung benutzt. GEIN verwaltet demnach ebenfalls Geoinformationen und ist in dieser Hinsicht ein raumbezogenes Metainformationssystem. Allerdings ist die räumliche Komponente der in GEIN verwalteten Informationen oftmals nicht das, was man herkömmlich davon erwarten würde. Viele der von GEIN referenzierten Ressourcen sind

¹⁶ German Enviromental Information Network

¹⁷ zu finden unter <http://www.gein.de/>

Dokumente, deren Raumbezug durch die Lokalität des beschriebenen Ereignisses erzeugt wird. Beispielsweise bekommt eine in Berlin gegebene Presseerklärung den Raumbezug Berlin und ist dann auch hierüber räumlich recherchierbar.

GEIN erlaubt im Gegensatz zum UDK den direkten Zugriff auf die Originaldaten, falls diese über eine Webseite verfügbar sind. GEIN selbst speichert allerdings nur Informationen über referenzierte Datensätze, wie Schlagworte und Referenzen dieser Daten. Allerdings integriert GEIN die verschiedenen Umweltinformationsressourcen, die unabhängig von GEIN bestehen, nur über spezielle Suchkriterien, während der UDK umfangreichere Informationen über die von ihm katalogisierten Informationsressourcen enthält.

GEIN ist daher ebenfalls ein Datenkatalog, wenn auch von sehr spezieller Ausprägung. Die von GEIN verwaltete Menge von Metadaten ist sehr gering und erlaubt nur auf sehr eingeschränkte Weise eine räumliche Recherche. Allerdings ist es möglich, auf Daten zuzugreifen, falls diese durch einen einfachen Link angesprochen werden können.

4.5.3 EEA, EIONET & CDS

Auf europäischer Ebene ist die *European Environmental Agency* (EEA) mit dem *European Environment Information and Observation Network* (EIONET) aktiv geworden. Das EIONET umfasst bislang Institutionen der 15 EU-Staaten und zusätzlich Island, Lichtenstein und Norwegen. Das EIONET besteht unter anderem aus europäischen Themenzentren, die bestimmte Aufgaben innerhalb des EIONET übertragen bekommen.

Eines dieser Zentren ist das beim Niedersächsischen Umweltministerium angesiedelte *European Topic Centre on Catalog of Data Sources* (ETC/CDS). Die Aufgabe des ETC/CDS ist die Konzeption und Entwicklung des *Directory of Information Resources* (DIR), welches einen Katalog von Umweltinformationsressourcen in Europa darstellt. Das DIR baut auf dem CDS, dem europäischen Metainformationssystem für Umweltinformationen auf. Das Datenmodell des CDS wiederum basiert auf dem Datenmodell des *Global Environmental Locator System* (GELOS) (LETHEN & JENSEN 1999).

Zur Verschlagwortung wird innerhalb des CDS der *General Multilingual Environmental Thesaurus* (GEMET) benutzt. GEMET ist ein mehrsprachiger, polyhierarchischer Thesaurus, der entwickelt wurde, um die Werkzeuge des CDS zu unterstützen. Er erlaubt einheitliches Katalogisieren und Indexieren von Umweltinformationen in ganz Europa (LETHEN 1999).

Der CDS ist ein Katalog für Umweltdaten in Europa und basiert auf dem UDK. Sein Ziel ist die Beantwortung von Fragen der folgenden Art: „WER ist für WELCHE Information in Europa verantwortlich, in WELCHER Form und WO liegen diese Daten vor, und WIE kann man sie finden“ (LETHEN 1999, S.1). Genau wie der UDK unterstützt der CDS einen *Discovery Service*, jedoch im Gegensatz zu diesem ist das gesamte Europa als Abdeckungsbereich anvisiert. Eine weitere Parallele zum UDK besteht darin, dass der CDS in einer Windows-Version, dem WinCDS, verfügbar ist, welcher die lokale Eingabe von Metadaten unterstützt. Darüber hinaus existiert eine internetgestützte Version¹⁸, welche die Recherche auf die zentral registrierten und verwalteten Metadaten ermöglicht. Der CDS erlaubt die Angabe eines minimal umschließenden Rechteckes für die von ihm referenzierten Informationsressourcen, so dass auch hier eine Georeferenzierung möglich ist.

4.6 Geoinformationen

Für die Definition raumbezogener Metadaten gibt es eine Vielzahl von Standardisierungsinitiativen, die zum großen Teil in die Entwicklung des bereits vorgestellten ISO-Standards mit Namen *Geographic Information – Metadata* einfließen. Drei dieser Standards, die relativ weite Verbreitung gefunden haben, sollen hier Erwähnung finden. Darüber hinaus werden einige inhaltlich interessante oder für die Metainformationsverarbeitung in Deutschland wichtige Initiativen Inhalt der folgenden Abschnitte sein. Der für Deutschland in der Zukunft sehr wichtige Metadatenstandard für das Amtliche Liegenschaftskataster Informationssystem (ALKIS), den Nachfolger des Amtlich-Topographischen Informationssystem (ATKIS) und der Amtlichen

¹⁸ zu finden unter <http://www.mu.niedersachsen.de/system/cds/>

Liegenschaftskarte (ALK), basiert auf ISO 19115 und wird deshalb nicht ausdrücklich diskutiert.

4.6.1 Content Standard for Digital Geospatial Metadata

Der *Content Standard for Digital Geospatial Metadata* (CSDGM) des *Federal Geographic Data Committee* (FGDC) entstand im Kontext der Aktivitäten der US-Regierung eine nationale Infrastruktur für Geodaten¹⁹ aufzubauen. Das FGDC setzt sich aus 17 Bundesbehörden der USA zusammen. Mittlerweile liegt der CSDGM in seiner zweiten Version vor. Eine Erweiterung für Fernerkundungsdaten steht kurz vor der Fertigstellung. Der Standard wurde aus der Sicht eines potenziellen Nutzers eines Geodatensatzes entwickelt, dem bei der Beurteilung des Datensatzes ein Anhaltspunkt gegeben werden soll. Er ist ein für jedermann verfügbares Dokument, welches die Metadatenelemente, ihre Namen, Datentypen und Kurznamen auflistet. Die Metadatenelemente werden, ebenso wie die von ISO 19115, unterteilt in verpflichtende, von einer bestimmten Bedingung abhängige und optionale Elemente. Verpflichtende Elemente müssen für jeden beschriebenen Datensatz angelegt werden, Elemente, die verpflichtend, wenn möglich sind, müssen angegeben werden, falls der beschriebene Datensatz ein bestimmtes Kriterium erfüllt. Im Falle von optionalen Elementen, ist es dem Beschreibenden freigestellt ist, ob er das Element angibt.

Der CSDGM, der auch oft als FGDC-Standard bezeichnet wird, beschreibt die Terminologie und Definition für die Dokumentation von digitalen Geodaten, unterteilt in zehn Kategorien. Diese sind (FGDC 1998):

- Identifikation
- Datenqualität
- Datenorganisation
- Raumbezug
- Entitäten und Attribute
- Datenabgabe
- Metadatenreferenz

¹⁹ *National Spatial Data Infrastructure* (NSDI)

- Zitierangaben
- Zeitliche Angaben
- Kontaktinformationen

Die ersten sieben dieser Kategorien geben die eigentliche Unterteilung der Metadaten vor und müssen somit für jeden zu beschreibenden Datensatz angelegt werden. Die letzten drei Kategorien beschreiben, wie Informationen zur Zitierweise, zeitbezogene Informationen und Kontaktangaben angelegt werden sollen, und werden innerhalb der anderen Kategorien benutzt. Die hier erwähnten zehn Kategorien entsprechen den Paketen von ISO 19115, wie sie in Abb. 5. gezeigt wurden, beziehungsweise stellen die Grundlage dieser Paketstruktur dar.

Der CSDGM stellt bis heute den international am weitesten verbreiteten Standard für raumbezogene Metadaten dar. Er war das Grundlegendokument von ISO 19115 und wird durch diesen Standard in der nahen Zukunft als Standardreferenz für Geo-Metadaten abgelöst werden.

4.6.2 Spatial Data Transfer Standard

Eng verknüpft mit dem CSDGM ist der *Spatial Data Transfer Standard* (SDTS) des FGDC. Das Ziel des SDTS ist der Datentransfer von raumbezogenen Daten mittels selbstbeschreibender Austauschdateien (ANSI 1997a). Eine solche Datei enthält räumliche Daten, Attribute, Informationen zur Georeferenzierung und Datenqualität, ein *Data Dictionary*, das technische sowie weitere notwendige Metadaten enthält. Mittels des SDTS soll es möglich sein, Geodaten zwischen verschiedenen Systemen verlustfrei auszutauschen. Der SDTS war eines der ersten Resultate der Anstrengungen der USA hinsichtlich des Aufbaus einer NSDI (GÜNTHER & VOISARD 1997).

Der SDTS beschreibt eine Möglichkeit zum Austausch von Geodaten mittels Dokumentation des Datenmodells auf konzeptioneller, logischer und physischer Ebene (ARCTUR ET AL. 1998). SDTS besteht aus drei Basisspezifikationen. Diese sind die logische Spezifikation, die Spezifikation räumlicher Objekte und eine auf ISO 8211 aufbauende *Encoding* Spezifikation. Die drei Teile beziehen sich zwar aufeinander, sind aber relativ unabhängig voneinander, da sie sich auf einzelne Teilaspekte der

Datenaustauschproblematik beziehen. Darüber hinaus existieren drei SDTS-Profile, die Regeln zum Datenaustausch spezifischer Datentypen definieren. Der erste Teil der Basisspezifikation, die *Logical Specifications* (ANSI 1997a) definiert das konzeptionelle Modell des SDTS, die räumlichen Objekttypen und die Beschreibungen der Datenqualität. Die hier verwendeten Metadaten überschneiden sich stark mit den Elementen des CSDGM. Der zweite Teil, mit Namen *Spatial Features* (ANSI 1997b) enthält eine Liste von *Entity Types*, die zur semantischen Kategorisierung räumlicher Objekte dienen. Im dritten Teil der Basisspezifikation, der den Namen *ISO 8211 Encoding* trägt (ANSI 1997c) wird beschrieben, wie SDTS-Dateien auf physischer Ebene gespeichert werden können.

Mit Hilfe des SDTS ist es möglich, Geodatensätze fast beliebiger Komplexität abzubilden und zu kodieren. Zusätzlich enthält ein SDTS-Datensatz alle zur Nutzung benötigten Metadaten. Er setzt sich somit aus Originaldaten, Informationen, welche die Datenstruktur beschreiben, und Informationen zum Entstehungskontext zusammen.

4.6.3 ENV Geographic Information – Data description – Metadata

Auf europäischer Ebene entstand im Rahmen des Technical Committee 287 des *Comité Européen de Normalisation* (CEN/TC 287) ein europäischer Prä-Standard mit dem Namen *ENV 12657 Geographic Information – Data description – Metadata*. CEN/TC 287 ist inzwischen inaktiv geworden, die letzte Sitzung fand Ende 1998 statt. ENV 12657 definiert ein konzeptionelles Schema für raumbezogene Datensätze. Das Schema sieht Informationen über den Inhalt, die zeitliche und räumliche Ausdehnung eines Datensatzes, das Raumbezugssystem, die Datenqualität und die Verwaltung eines Datensatzes vor. ENV 12657 identifiziert genau wie ISO 19115 und der CSDGM einen Satz von Kernelementen, die für jeden beschriebenen Datensatz anzugeben sind. Es werden Beispiele gegeben, wie der Standard anzuwenden ist, allerdings finden sich keine Hinweise auf die konkrete Datenbankumsetzung oder ein Datenbankschema (CEN 1998).

Die Ergebnisse von ENV 12657 gingen genau wie die vieler anderer Standards in die Entwicklung von ISO 19115 *Geographic Information – Metadata* mit ein.

4.6.4 InGeo-MIS und InGeo-MDF

Das InGeo-Forum ist ein in Darmstadt ansässiger und 1997 gegründeter Verbund verschiedener Institutionen, die sich mit Geoinformation befassen. Ziel des InGeo-Forums ist die Stimulierung des Geodatenmarktes und die Verfügbarmachung von Geodaten für die Öffentlichkeit. Es sieht sich als Informations- und Kooperationsforum zwischen Datenanbietern und Datennutzern (INGEO-FORUM 2000). Eine der Initiativen zur Erreichung des Zieles, eine nationale Geodateninfrastruktur zu forcieren, ist der Aufbau des InGeo-Metainformationssystems, das bislang prototypisch implementiert ist und im Laufe des Jahres 2002 in den produktiven Einsatz gehen soll²⁰. Das InGeo-MIS ist ein Internet-basiertes Metainformationssystem für Geodaten, das mittels textueller und räumlicher Recherchekriterien die Suche nach Geodaten erlaubt. Es besteht die Möglichkeit, andere Metainformationssysteme in das InGeo-MIS einzuhängen, was prototypisch für den virtuellen UDK bereits durchgeführt wurde. Zu Beginn des Aufbaus des InGeo-MIS existierte der ISO Standard 19115 nur in der ersten Entwurfsversion²¹. Da diese Version noch nicht sehr stabil und außerdem nicht völlig ausgereift schien, wurde für das InGeo-MIS ein eigenes, wenn auch auf ISO 19115 aufbauendes Metadatenformat geschaffen, das InGeo-Metadatenformat (InGeo-MDF, BALFANZ & GÖBEL 1999). Zwischenzeitlich wurde dieses Metadatenformat überarbeitet und teilweise an die zweite Entwurfsversion von ISO 19115 angepasst. Es liegt derzeit in der Version 2.1 vor. Das InGeo-MDF stellt ein *Profile* des ISO-Standards dar. Es wurde eine große Menge von ISO-Elementen in das InGeo-MDF aufgenommen und einige Erweiterungen eingeführt. Die Anzahl der Erweiterungen ist allerdings eher gering und sie überschneiden sich teilweise mit den Neuerungen, die am ISO-Standard durchgeführt wurden.

²⁰ zu finden unter <http://www.ingeo-ic.de>

²¹ Zu diesem Zeitpunkt war ISO 19115 noch unter dem Namen ISO 15046-15 bekannt.

Das InGeo-MIS stellt einen Geodatenkatalog im Sinne des OGC dar, auch wenn keine der OGC-Spezifikationen beim Entwicklungsprozess eine Rolle spielten. Es erlaubt die Recherche (*Discovery*), den direkten Datenzugriff (*Access*) und das Management von Geodaten und Metadaten.

4.6.5 GIS-Pool

GIS-Pool ist der Name einer Managementkomponente für Geodaten, die für das Umweltministerium von Sachsen Anhalt entwickelt wurde (FITZKE ET AL. 2000). Das GIS-Pool Konzept sieht vor, dass Anbieter von Daten ihre verteilten Geodatenbestände bei einem Datenkatalog anmelden und potenzielle Nutzer der Daten über ein differenziertes Rechtekonzept auf diese Daten zugreifen können. Dies entspricht einem *OpenGIS* Geodatenkatalog, der *Discovery*, *Access* und *Management* Dienste zur Verfügung stellt.

GIS-Pool benutzt einen Metadatenkatalog, der auf ISO 19115 aufbaut und mittels einer Im- und Exportschnittstelle auf das UDK-Metadatenmodell abgebildet werden kann. Zusätzlich zu der textbasierten Metainformation wird hier allerdings dem Nutzer ein spezielles Vorschaubild angezeigt, welches in Anlehnung an den aus der Fernerkundung stammenden Begriff im Rahmen des Projektes *Footprint* genannt wird. Dieser *Footprint* beinhaltet Informationen über räumliche Verteilung der vorliegenden Geodaten als auch Informationen über die Aktualität der Daten mit Hilfe einer Rasterung der dargestellten Fläche sowie unterschiedlicher Farbwahl der einzelnen Rasterzellen (GREVE, LESSING & MÜLLER 2001). Dieser Ansatz ist insofern von besonderem Interesse, da Metainformation aus einer Verbindung von einem einfachen minimal umgebenden Rechteck zusammen mit Informationen, die aus den eigentlichen Datensätzen abgeleitet werden, erzeugt wird. Ein solches Vorschaubild kann durch Variierung der Rasterweite beliebig genaue Informationen über den eigentlichen Datensatz geben, was dazu führt, dass Originaldaten dynamisch zur Erstellung von Metadaten benutzt werden. Die strenge Unterscheidung zwischen Metadaten und Originaldaten ist hierdurch nicht mehr durchführbar, sondern kann nur mit Hilfe des Benutzungskontextes getroffen werden.

4.6.6 Geodateninfrastruktur Nordrhein-Westfalen

Die Geodateninfrastruktur Nordrhein-Westfalen (GDI NRW) ist eine vom Land NRW ins Leben gerufene Initiative zur Aktivierung des Geodatenmarktes. Die Situation im Hinblick auf die aktuelle Nutzung von Geodaten, deren Potenziale sowie mögliche Handlungsschritte wurde in einer Marktstudie festgehalten (FORNEFELD & OEFINGER 2001). Diese Marktstudie fasst Motivation und Ziele der GDI aus wirtschaftlicher Sicht zusammen.

GDI NRW ist im Gegensatz zu den anderen hier vorgestellten Informationssystemen kein zentral verwalteter Geodatenkatalog, sondern entsteht durch Zusammenarbeit von Softwarefirmen und Datenanbietern, die monetär abrechenbare, raumbezogene Dienste zur Verfügung stellen. Die Harmonisierung und Zusammenarbeit der verschiedenen Dienste wird mittels eines *Testbeds* hergestellt. Diese *Testbeds* stellen Testläufe dar, bei denen von den Beteiligten verteilte Dienste zur Verfügung gestellt werden, die dann von den anderen Teilnehmern benutzt werden können. Um diese Dienste recherchierbar zu machen, werden *Web-Registries* benutzt, Kataloge, die auf Dienste verweisen. Diese Dienstorientiertheit von GDI NRW hat zur Folge, dass in Hinsicht auf Metadaten der Standard ISO 19119 eine große Rolle spielt und erst in zweiter Hinsicht ISO 19115.

GDI NRW ist eine eng an das OGC gekoppelte Initiative, die Architektur der GDI ist vollständig auf OGC-Spezifikationen aufgebaut. Das Konzept der Katalogdienste ist von zentraler Bedeutung und wird hier konsequent auf raumbezogene Dienste angewandt.

4.6.7 GeoMIS.Bund

Das deutsche Regierungskabinett hat 1998 den Beschluss gefasst, einen Interministeriellen Arbeitskreis für Geoinformation (IMAGI) einzurichten, dessen Vorsitz und Geschäftsführung das Bundesministerium des Inneren inne hat. Ziel des IMAGI ist es, den Zugang zu Geoinformationen durch bessere Koordinierung und Einsatz moderner Technologie zu verbessern.

Ein erster Schritt zur Erreichung dieser Ziele ist der Aufbau eines Metainformationssystems für Geodaten des Bundes, das einen einheitlichen Zugang zu den Metadaten der Bundesverwaltungen schaffen soll. Dieses

GeoMIS.Bund übernimmt eine Brokerrolle, die es erlauben soll, mittels eines Portales mit Hilfe verteilter Anfragen an die über 50 bereits existierenden Fachmetainformationssysteme Datensätze zu identifizieren und Kontaktinformationen zu erhalten.

Das Datenmodell des GeoMIS.Bund beruht auf ISO 19115 (IMAGI 2000). Die Konzeption von GeoMIS.Bund sieht vor, dass die einzelnen Fachmetainformationssysteme als autarke Geodatenkataloge behandelt werden, die mittels der vom OGC spezifizierten Katalogschnittstelle (OGC 1999e) angesprochen werden. Bislang ist allerdings nur der Zugriff auf die Metadaten vorgesehen, die *Discovery*-Schnittstelle der *Catalog Interface Implementation Specification* wird als einzige implementiert (siehe auch Abb. 8.). Das Portal von GeoMIS.Bund stellt selbst auch einen Katalogdienst dar und kann somit ebenfalls über eine solche Schnittstelle angesprochen werden. Diese Konzeption ergibt somit einen kaskadierenden Katalogdienst, der auf andere Katalogdienste zur Beantwortung von Anfragen zugreift. Eine eigene Metadatenhaltung des zentralen Katalogdienstes ist nicht vorgesehen.

4.7 Digitale Bibliotheken

Digitale Bibliotheken stellen die Weiterentwicklung von traditionellen Bibliotheken dar, wie sie von Bibliothekaren seit Jahrhunderten verwaltet werden. Der Begriff der digitalen Bibliothek wird oftmals auch freier interpretiert, um zum Beispiel Sammlungen von Algorithmen, Online-Datenbanken oder das gesamte *World Wide Web* zu bezeichnen. In der erstgenannten engeren Definition besitzen Digitale Bibliotheken die folgenden Eigenschaften (CLEVELAND 1998):

- Sie sind die digitale Form traditioneller Bibliotheken, die sowohl digitale als auch Papiermedien umfassen.
- Sie umfassen zusätzliches, digitales Material, das außerhalb der physikalischen und verwaltungstechnischen Grenze jeder digitalen Bibliothek existiert.
- Digitale Bibliotheken enthalten alle Vorgänge und Dienste, die das Nervensystem von Bibliotheken darstellen. Traditionelle Vorgänge

müssen allerdings wegen der Unterschiede zwischen neuen und alten Medien überarbeitet werden.

- Sie sollten im Idealfall eine einheitliche Sicht auf alle Information, die innerhalb einer Bibliothek enthalten ist, zur Verfügung stellen und zwar unabhängig von Form und Format.
- Sie dienen bestimmten Gemeinschaften oder Institutionen auf gleiche Weise, wie das bislang traditionelle Bibliotheken tun, obwohl diese Gemeinschaften verteilt im Netz existieren können.
- Sie setzen die Fertigkeiten des Bibliothekars als auch des Informatikers voraus.

Ausgehend von diesem erweiterten Bibliotheksbegriff wird die Tatsache offensichtlich, dass Metadaten eine zentrale Stellung im Problembereich digitaler Bibliotheken einnehmen, denn die Erfassung und Verwaltung von Metadaten ist die Hauptaufgabe von Bibliothekaren. Die *Anglo-American Cataloguing Rules* sind ein nationales Beispiel für die Standardisierung von Metadaten im Bibliotheksbereich. Da diese spezialisierten Metadaten allerdings extrem umfangreich und zeitaufwändig zu benutzen sind, bestand Bedarf an einfacheren Standards. Das *Dublin Core Element Set* (DCES) ist aus dieser Anforderung heraus entstanden. Das DCES ist in einer Folge von acht Workshops entwickelt worden, an denen ein internationales multidisziplinäres Team von Experten teilnahm. Ergebnis ist ein Satz von 15 Metadatenelementen, die es erlauben, Informationsressourcen zu beschreiben. Diese Metadatenelemente sollen allerdings nur einen Kern an standardisierten Metadaten zur Verfügung stellen, erheben also keinen Anspruch auf Vollständigkeit. Die ursprüngliche Intention der Nutzung der DCES war die Ausrichtung auf Dokumente. Durch die Allgemeinheit der Elemente können jedoch auch viele andere Arten von digitalen Ressourcen beschrieben werden (DCMI 2001).

4.8 Internet

Die fehlende Strukturierung von Internetressourcen führt dazu, dass es schwer ist, im *World Wide Web* zu einer konkreten Fragestellung die passende Antwort zu finden. Lösungen für dieses Problem werden von Suchmaschinen

und Internetportalen angeboten, die Internetressourcen auf manuelle, semiautomatische oder automatische Weise katalogisieren. Um diese Katalogisierung in geregelte Bahnen zu lenken, ist Metainformation notwendig, die sowohl den Datenanbietern erlaubt, ihre Informationsangebote für die von ihnen beabsichtigten Nutzergruppen verfügbar zu machen, als auch den Nutzern die Recherche vereinfacht. Zu diesem Zweck enthält ein HTML-Dokument die *Meta-Tags*²², welche die Beschreibung von Webseiten erlauben. Zwischenzeitlich durchsuchen Agenten und Roboter, die Suchmaschinen zuarbeiten, nicht nur diese speziellen Tags, sondern den gesamten Inhalt von Webseiten nach Schlüsselwörtern.

XML ist ursprünglich dazu entwickelt worden, Informationen des Internets strukturiert zu speichern und recherchierbar zu machen. Inzwischen hat sich diese Technologie allerdings in vielen anderen Bereichen etabliert und wird weitverbreitet als Austauschformat für Informationen aller Art benutzt. XML ist aus verschiedenen Gründen zur Metainformationsverarbeitung sehr gut geeignet. XML ist inzwischen nicht mehr nur eine Datenauszeichnungssprache, sondern wurde durch verschiedene XML-Anwendungen erweitert, die den Umgang mit dieser Technologie im Metadatenumfeld erst attraktiv machen. Dies sind insbesondere XML-Schema, XML-RDF und XSLT, die in den nächsten Abschnitten Erwähnung finden sollen.

XML ist eine Sprache zur Definition von Sprachen, oder genauer eine Auszeichnungssprache für Dokumente, die strukturierte Information enthalten. Sie ist somit eine Meta-Sprache und konzeptionell keinesfalls auf Ebene der *Hypertext Markup Language* (HTML) anzusiedeln. Mit Hilfe von XML ist es möglich, Sprachen wie HTML zu definieren²³. XML ist eine Teilmenge der *Standard Generalized Markup Language* (SGML), aber im Gegensatz zu dieser bedeutend einfacher zu verstehen und zu implementieren. Dies ist der Hauptgrund dafür, dass sie sich in allen Anwendungen, die mit dem Internet in Zusammenhang stehen, in erstaunlich kurzer Zeit verbreitet hat²⁴. Die

²² Auszeichnen

²³ XHTML, siehe <http://www.w3.org/TR/2000/REC-xhtml1-20000126/>

²⁴ der erste XML-Entwurf wurde im November 1996 veröffentlicht

Benutzung von XML zur Ablage von Metadaten ist so naheliegend wie praktisch, da XML-Dokumente direkt lesbar sind und keine Formatierungsinformationen enthalten. Innerhalb eines XML-Dokumentes wird nur die Struktur und der Inhalt der Daten mit Hilfe von *Tags* definiert. Bei intelligenter Definition dieser *Tags*, das heißt, wenn sie mit semantisch sinnvollen Namen belegt werden, ist es möglich, ein XML-Dokument mit Hilfe eines herkömmlichen Editors und ohne weiteres Vorwissen zu lesen. Durch die Trennung von Inhalt und Darstellung ist darüber hinaus die Möglichkeit gegeben, eine XML-Datei auf unterschiedliche Weisen darzustellen. Man kann davon ausgehen, dass in absehbarer Zeit jegliche digital gespeicherte Metainformation mit Hilfe von XML verwaltet wird oder zumindest als Austauschformat herangezogen wird.

Es existiert eine Reihe von Co-Standards, die auf unterschiedliche Weise mit XML in Verbindung stehen. Für die Metainformationsverarbeitung sind drei dieser Standards von besonderer Bedeutung: XML Schema, RDF und XSLT. XML Schema ist eine Möglichkeit, XML-Dokumente zu beschreiben und einzuschränken. Sie stellen somit eine Alternative zu *Document Type Definitions* (DTD) dar (VLIST 2000). Mit Hilfe solcher Einschränkungen ist es möglich, spezielle Sprachen zu definieren, indem Struktur und Attributdefinition vorgegeben wird. XML Schema erlaubt im Gegensatz zu DTDs zusätzlich die Definition von Wertebereichen und Vererbung, Funktionalitäten, die für die Standardisierung von Metadaten von großer Wichtigkeit sind.

Das *Resource Description Framework* (RDF) definiert einen Rahmen für Beschreibung und Austausch von Metadaten mit Hilfe von XML. Der Anwendungskontext für RDF ist das Internet, es bietet folglich Möglichkeiten zur Beschreibung jeglicher Ressource, die eine Web-Adresse besitzt. RDF benutzt XML als Austauschformat. RDF besitzt zwei Unterschiede gegenüber gewöhnlichem XML:

- die Reihenfolge der Elemente innerhalb eines Dokumentes ist nicht relevant und
- es besitzt eine eingeschränkte Datenstruktur, so dass RDF-Dokumente leichter von Computern in großer Menge zu verwalten sind.

XML-RDF stellt eine Technologie dar, welche die Meta-Tags von HTML ablösen sollte, die aber bislang zu diesem Zweck kaum umgesetzt wird.

Die *eXtensible Stylesheet Language Transformations* (XSLT) ermöglichen aber erst die Interoperabilität, die von XML versprochen wird, denn hiermit können in XML kodierte Daten von einem Schema (definiert durch DTD oder XML-Schema) in ein anderes überführt werden (HOLMAN 2000). Im bereits angesprochenen GIS-Pool Projekt wurde XSLT beispielsweise verwendet, um eine UDK-Schnittstelle zu implementieren. Dies geschieht mit Hilfe einer Transformation, die die auf ISO 19115 aufbauenden Metadatenelemente des Kataloges auf die entsprechenden UDK-Elemente abbildet.

Abgesehen von diesen eher technischen Entwicklungen hat sich das bereits erwähnte DCES inzwischen als de facto Standard für die Beschreibung von allgemeinen Informationsquellen im Internet etabliert (PICK & SAARENMA 2001). Die Umsetzung der DCES ist beispielsweise mittels normalem XML oder RDF möglich (PICK, KVIST & LOUNMA 2001).

Die neueste Entwicklung im Bereich Metadatenrecherche im Internet ist das *Open Archives Initiative* (OAI), die ein Protokoll zur interoperablen Recherche in Metadatenarchiven entwickelt hat (SOMPEL & LAGOZE 2001). Dieses auf HTTP²⁵ aufbauende Protokoll benutzt wiederum die DCES zur Definition von Metadatenelementen.

4.9 Medizininformatik

Das übliche Medium zur Kommunikation medizinischer Information ist der medizinische Bericht. Medizinische Berichte oder andere medizinische Dokumente bestehen aus textbasierten Daten wie Beschreibungen von Patienten, Untersuchungen oder Symptomen und darüber hinaus multimedialen Daten wie zum Beispiel Röntgenbildern. Medizinische Dokumente sind deshalb als Multimedia-Dokumente zu betrachten.

Moderne Systeme der Gesundheitsvorsorge müssen sich mit Problemen der effizienten Verwaltung von multimedialen Dokumenten und mit dem Problem der Integration von verteilten Daten auseinandersetzen, da sie oft aus verteilten

²⁵ Hyper Text Transfer Protocol

Strukturen bestehen, die ihre Informationen in autonomen und heterogenen Umgebungen verwalten (CONSORTI ET AL. 1996). Für beide Probleme ist der Einsatz von Metadaten eine naheliegende Lösungsmöglichkeit. Aus diesem Grunde haben sich einige spezielle Ansätze zum Management von medizinischer Information mittels Metadaten gebildet. Das Hauptanliegen der Verfügbarmachung von medizinischen Informationen für verteilte Benutzergruppen hat zu einer Konzentration der Metadatenaktivitäten auf den Internet-Bereich geführt und zur Nutzung der DCEs auch für medizinische Dokumente.

4.10 Zusammenfassung

Ein Überblick der verschiedenen Kontexte, in denen Metainformationsverarbeitung betrieben wird, zeigt sowohl die Unterschiedlichkeit der Ansätze als auch Gemeinsamkeiten (siehe Tab. 1). Zusätzlich zu den in diesem Kapitel vorgestellten Initiativen wird in der Tabelle der schon früher vorgestellt ISO Standard 19115 mit aufgeführt.

Im Umfeld von Datenbanken werden rein technische Metadaten in *Datenbank-Dictionaries* verwaltet, Entstehungskontext und andere semantische Information hat hier keine Bedeutung. Die hier verwalteten Metadaten werden vom Datenbankmanagementsystem zur Datenverwaltung benötigt. Diese Metadaten beschreiben die Struktur eines Datenbankschemas. Obwohl sie in erster Hinsicht zur Verarbeitung durch eine Maschine gedacht sind, können sie auch von Menschen gelesen werden, die eine Katalogsicht auf vorhandene Datenbestände benutzen wollen.

Die Erweiterung dieses Ansatzes im Umfeld von *Data Warehouses* bezieht semantische Metainformation mit ein. Der Grund hierfür liegt darin, dass die Integration von Daten verschiedener Quellen die Hauptaufgabe eines *Data Warehouses* ist. Zur Integration dieser Datenquellen ist es notwendig, sehr umfangreiche Informationen über deren Herkunft und Entstehung zu besitzen. Die Beschreibung des Kontextes, aus welchem die Daten stammen, ist hier von großer Bedeutung.

Tabelle 1. Überblick über genannte Metadaten-Initiativen

Name	Typ	Wirkungsbereich	Anwendung
UDK	Informationssystem	Deutschland und Österreich	Katalogisierung von Umweltdaten der öffentlichen Verwaltung
GEIN	Portal	Deutschland	Indexierung umweltrelevanter Internetressourcen
CDS	Informationssystem	Europa	Katalogisierung von Umweltinformationen
CSDGM	Standard	International (vor allem USA)	Standardisierung räumlicher Metadaten
SDTS	Austauschformat	International (vor allem USA)	Austausch von Geodaten
ENV 12657	Standard	Europa	Standardisierung räumlicher Metadaten
ISO 19115	Standard	International	Standardisierung räumlicher Metadaten
InGeo-MIS	Informationssystem	Deutschland	Katalogisierung von Geodaten
InGeo-MDF	Standard	Deutschland	Auf ISO 19115 aufbauende Standardisierung räumlicher Metadaten
GIS-Pool	Informationssystem	Umweltministerium Sachsen-Anhalt	Management von Metadaten in verteilten Institutionen
GDI NRW	Infrastrukturprojekt	Nordrhein-Westfalen	Wertschöpfung von Geodaten mittels verteilter Dienste
GeoMIS.Bund	Informationssystem	Deutschland	Katalogisierung von Geodaten der öffentlichen Verwaltung
DCES	Standard	International	Metadaten für beliebige Ressourcen, vor allem Dokumente
OAI	Protokoll	International	Interoperable Recherche in Metadatenarchiven

In verschiedenen Bereichen, die mit der Dokumentation und Integration unterschiedlich komplexer Dokumente befasst sind, hat sich das *Dublin Core Element Set* weitgehend durchgesetzt. Ursprünglich aus dem Bibliothekswesen stammend, hat sich dieser Minimalstandard zur Beschreibung von Informationsressourcen im Internet, dem Umweltinformationswesen, der medizinischen Informatik und in anderen Bereichen etabliert. In diesen Anwendungen ist die semantische Dokumentation von hauptsächlicher Bedeutung, während technische Metadaten nur eine untergeordnete Rolle

spielen. Das DCES stellt eine Möglichkeit dar, Informationsressourcen zur Verwendung einer Katalogisierung zu beschreiben.

In der raumbezogenen Metainformationsverarbeitung wird sowohl technische als auch semantische Metainformation benötigt, um die umfangreichen Interoperabilitätsprobleme lösen zu können. Der Raumbezug von Informationen bedarf allerdings besonderer Beachtung, da hierdurch sowohl zusätzliche Integrationsprobleme als auch zusätzliche Möglichkeiten entstehen, Informationen zueinander in Bezug zu setzen. Die Erzeugung von Geodaten ist sowohl zeit- als auch kostenintensiv, weswegen ihre Mehrfachnutzung von großer Bedeutung ist. Das Hauptproblem ist hier immer noch die meist fehlende Dokumentation des Entstehungskontextes von Geodaten, die hierfür notwendig wäre. Darüber hinaus ist es notwendig, die vorhandenen Datenbestände für zahlende und nichtzahlende Zweitnutzer verfügbar zu machen. Metainformationssysteme beziehungsweise Katalogsysteme dienen diesem Zweck und benötigen hierzu ebenfalls Metainformationen, die zwar große Überschneidungen mit den kontextbeschreibenden Informationen haben, aber nicht vollständig identisch sind. Zur Katalogisierung sind oftmals weniger umfangreiche Metainformationen notwendig als zur Beschreibung des Entstehungskontextes benötigt werden. Allerdings gibt es auch Metainformationen, die speziell für Kataloge von Bedeutung sind. Vor allem sind dies Informationen zur Datenabgabe wie Zugriffswege und Preisinformationen. Darüber hinaus gewinnen Katalogsysteme stark an Attraktivität, wenn sie die Möglichkeit besitzen, Vorschauinformationen über die räumliche Lage ihrer Datenbestände zur Verfügung zu stellen. Solche Vorschaubilder werden aus den Originaldaten hergestellt. Sie sind somit Teildatensätze, werden aber in Katalogen mit der Funktion von Metadaten benutzt.

Die Verarbeitung von Geoinformationen erfolgt zunehmend mit Hilfe von Diensten, die einzelne Verarbeitungsaspekte kapseln und in komplexe Informationssysteme integrierbar sind. Aus diesem Grund gewinnen Metadaten, die nicht nur auf Geodatensätze anwendbar sind, sondern auch auf Geodienste, zunehmend an Bedeutung. Dienste beschreibende Metadaten werden in Katalogen aufgeführt, um diese Dienste für potenzielle Nutzer

identifizierbar zu machen und um den Zugriff auf sie zu beschreiben. Katalogsysteme im Sinne des OGC machen keinen Unterschied zwischen Daten referenzierenden und Dienste beschreibenden Katalogen.

Zwischen den verschiedenen Ansätzen bestehen vielfältige Verknüpfungen. Umweltinformationssysteme beispielsweise dokumentieren sowohl Dokumente als auch Geoinformationen. Sie werden teilweise auf Ansätzen des *Data Warehousing* aufgebaut und besitzen oftmals eine Auskunftskomponente im Internet. Geografische Informationssysteme haben Beziehungen zu verschiedenen anderen Technologien, unter anderem zu klassischen Datenbanken, und werden inzwischen weit verbreitet mit Internet-Auskunftsplätzen ausgestattet. Eine Analyse der Metainformationsverarbeitung kann somit nicht gesondert für die einzelnen aufgeführten Bereiche durchgeführt werden.

Aufbauend auf den in diesem Kapitel vorgestellten Ansätzen zur Metainformationsverarbeitung ist es möglich, verschiedene Funktionen von Metainformationen zu unterscheiden. Diese Unterscheidung ist nicht auf raumbezogene Metainformationen beschränkt, sondern kann auf alle Arten angewandt werden. Grundsätzlich kann in allen Ansätzen zwischen drei Aufgaben unterschieden werden:

- Dokumentation technischer Zusammenhänge zur Verwaltung und Konsistenzsicherung von Datenbeständen,
- Dokumentation des Datenkontextes zur Integration und Wiederverwendung von Daten und
- Katalogisierung von Datenbeständen und Diensten für Auffindung und Zugriff.

Diese Unterscheidung dient als Grundlage der im nächsten Kapitel durchgeführten Modellbildung der raumbezogenen Informationsverarbeitung.

5 Ein Modell der raumbezogenen Metainformationsverarbeitung

5.1 Beschreibung des Ansatzes

Wie im letzten Kapitel gezeigt, geschieht Definition und Umgang mit (raumbezogener) Metainformation bislang intuitiv und uneinheitlich. Zur Lösung dieses Problems wird als Ziel dieser Arbeit die Entwicklung eines allgemeinen Modells der raumbezogenen Metainformationsverarbeitung definiert. Diese Theorieentwicklung beruht vor allem auf Erkenntnissen der Semiotik, die sich im Allgemeinen mit Zeichen, deren Semantik und Interpretation befasst. Repräsentationen von Geodaten, das heißt sowohl kartografische Darstellungen, als auch die sie beschreibenden Metadatensätze, sind ebenfalls Zeichen und können deshalb mit Ansätzen aus der Semiotik untersucht und definiert werden.

Ein erster Entwurf dieses Modells wurde in MÜLLER (2001b) vorgestellt. Der dort vorgestellte Ansatz wird in dieser Arbeit weiterentwickelt.

5.2 Objektorientierte Denkweise

Zur Definition des theoretischen Unterbaus, der unbedingt notwendig ist, um das schwer zu greifende Problem der Metainformationsverarbeitung besser zu fassen, wird auf die objektorientierte Denkweise zurückgegriffen. Zur Darstellung der in der objektorientierten Modellierung verwendeten Konstrukte wird die *Unified Modelling Language* (UML) benutzt. Zu Definition und Umfang der UML sei auf die einschlägige Literatur verwiesen²⁶. Die in dieser Arbeit verwendeten Konstrukte werden im Anhang näher erläutert.

Die Verwendung der UML erlaubt es, komplexe Zusammenhänge zwischen den verschiedenen Klassen und Objekten der Metainformationsverarbeitung anschaulich darzustellen.

²⁶ beispielsweise OESTEREICH (1998)

5.3 Informationen und Daten

Es existiert eine Fülle von Definitionen für die Begriffe Information und Daten, die sich je nach Sichtweise und disziplinärem Hintergrund unterscheiden. Allen diesen Definitionen ist gemeinsam, dass Informationen etwas Höherwertiges darstellen als Daten. In der semiotischen Literatur wird der Begriff Information allerdings meist nicht direkt benutzt, sondern die Worte „Signifikat“, „Sinn“ oder „Begriff“ verwandt (ECO 1977). Alleine BERTIN (1974) gibt eine eindeutige Definition für diesen Begriff, der zudem die Definition grafischer Information unterstützt. Im Kontext der hier vorliegenden Arbeit wird deshalb in Anlehnung an BROY (1992) und vor allem BERTIN (1974) Information folgendermaßen definiert:

Information nennt man den abstrakten Gehalt („Bedeutungsinhalt“, „Semantik“) eines Gedankens. Dieser Gehalt besteht aus einer oder mehreren Beziehungen („Assoziationen“) zwischen einer Invariante (dem Begriffsinhalt, der allen Beziehungen gemeinsam ist) und einer Menge von Variationsbegriffen (auch: „Klassen“, „Komponenten“).

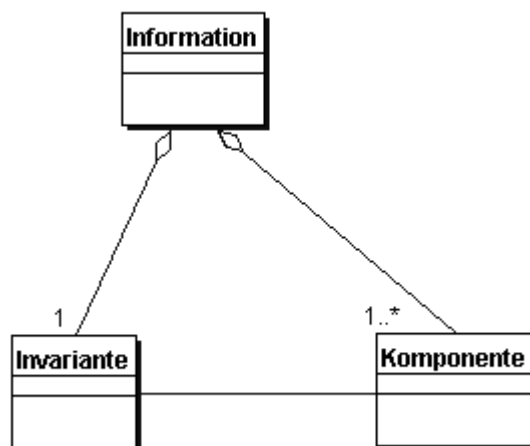


Abbildung 10. Information

Beispiel: „Blauwale sind migrierende Tiere“ ist ein möglicher Gedanke (eine Information). Die Semantik dieses Gedankens besteht aus der Invariante „Blauwal“ und der Klasse Tiergruppen. „Migrierende Tiere“ ist eine Kategorie dieser Gruppe.

Der Sinn von Informationen ist die Unterstützung von Entscheidungen (FRANK 1996, ARONOFF 1991). Informationen erlauben die Beurteilung von

Sachverhalten, was es Menschen erlaubt, (rationale) Entscheidungen zu treffen.

Zwischen Informationen und Daten besteht ein enger Zusammenhang. In Anlehnung an GREVE (1994) wird folgendes definiert:

Daten sind atomisierte Informationen. Sie kodieren die einzelnen Beziehungen, die eine Information ausmachen.

Beispiel: Die drei derzeit am häufigsten benutzten Möglichkeiten atomare Daten zu speichern sind (Text-)Dateien, relationale Datenbanken und objektorientierte Datenbanken. Innerhalb von Textdateien werden Daten durch ihre Zugehörigkeit in eine Zeile verknüpft, das gleiche gilt für relationale Datenbanktabellen, wobei in einer relationalen Datenbank die Tabelle (= Relation) grundsätzlich einen eigenen Namen besitzt und die Spalten ebenfalls mit Namen versehen werden müssen.

Tabelle 2. Zoologische Arten (Ausschnitt aus GROMS-Datenbank)

Latein	Englisch	Deutsch
Eubalaena australis	Southern right whale	Südkaper
Eubalaena glacialis	Northern right whale	Nordkaper
Eudromias morinellus	Eurasian dotterel	Mornellregenpfeifer
Falco naumanni	Lesser kestrel	Rötelfalke
Fulica atra	Common coot	Bläßhuhn
Gallinago gallinago	Common snipe	Bekassine

Tabelle 2 zeigt einen Ausschnitt aus der Datenbank von GROMS²⁷ wobei ein Teil der Arten-Relation dargestellt ist. Hier wären die Namen der Daten „Latein“, „Englisch“ und „Deutsch“ und die Werte die zugehörigen Felder in der Tabelle.

Objektorientiert gespeicherte Daten können in diesem Zusammenhang als Übermenge relationaler Daten gesehen werden. Innerhalb einer objektorientierten Datenbank würde ein Objekt der Klasse „Zoologische Art“

²⁷ *Global Register of Migratory Species*, zu finden unter www.groms.de

existieren, das die Attribute „Latein“, „Englisch“ und „Deutsch“ besitzen würde (Abb. 11.).



Abbildung 11. Diagramm der Klasse „Zoologische Art“

Die zoologische Art in ihrer Interpretation als Klasse umfasst somit den gesamten Informationsgehalt, den die Relation gleichen Namens umfasst, gibt aber darüber hinaus die Möglichkeit, Methoden zu definieren.

5.4 Zeichen und Kode

Der Zeichenbegriff wird hier im Sinne der Semiotik verwendet. Von ECO (1977, 1987) werden Zeichen folgendermaßen definiert:

Ein Zeichen ist die Korrelation eines Signifikanten (das Bezeichnende) mit einer Einheit (oder einer Hierarchie von Einheiten), die wir als Signifikat (das Bezeichnete) definieren. Oder einfacher: Zeichen sind etwas, was als Stellvertreter für etwas anderes steht.

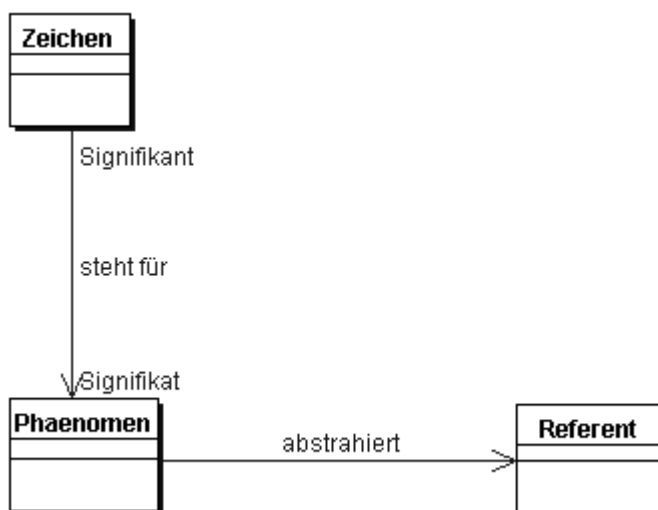


Abbildung 12. Zeichen

Die Beziehung zwischen einem Zeichen und dem bezeichneten Phänomen ist in Abb. 12. dargestellt. Der Zeichenprozess ist somit ein Kommunikationsprozess, bei dem Information von einem Sender an einen Empfänger weitergegeben werden soll. Bei jedem Zeichenprozess sind zu unterscheiden: das Signifikat, welches das eigentliche Zeichen ist, der Signifikant, welcher der bezeichneten Information entspricht und dem Referent, welcher dem bezeichneten Gegenstand entspricht (siehe Abb. 12.). Digitale Daten sind ebenfalls Zeichen, denn sie dienen als Darstellungsform für eine Information und sind somit Informationsträger (HAKE, GRÜNREICH & MENG 2001).

Beispiel: Die folgende Abbildung zeigt unter anderem fünf Bedienungsknöpfe einer Softwareoberfläche.

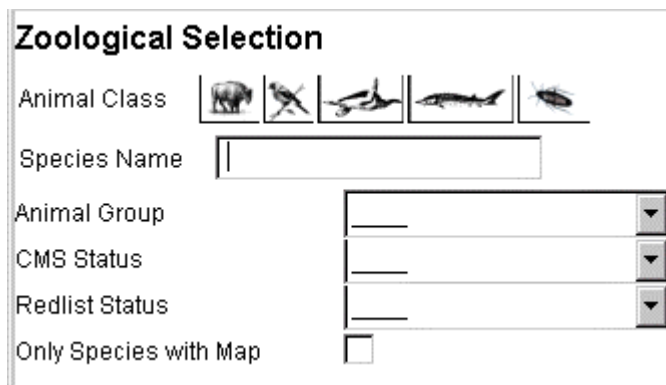


Abbildung 13. Ausschnitt aus der Benutzeroberfläche von GROMS

Diese fünf Bilder erfüllen die Funktion von ikonischen Zeichen (ECO 1987). Das Bild entspricht dem Signifikanten, der Referent ist beim Beispiel des Fisches ein Stör (Ordnung *Acipenseriformes*). Das Signifikat des Zeichens ist die Information, dass die zoologische Klasse *Pisces* (Fische) nach Drücken des Knopfes ausgewählt ist.

Angelehnt an ECO (1977) ist die folgende Definition:

Ein Kode stellt semantische Äquivalenzen zwischen den Elementen eines Signifikantensystems (dem Zeichen) und denen eines Signifikatsystems (dem Bezeichneten) her.

Ein Kode ist somit ebenfalls eine Beziehung zwischen einer Invariante (dem Signifikat) und einer Komponente (dem Signifikanten) und somit ebenfalls eine Information im Sinne der hier verwendeten Definition.

Beispiel: Der Kode im Falle der bei GROMS verwendeten ikonischen Zeichen ist nicht explizit angegeben. Es wird vorausgesetzt, dass die Assoziation zwischen dem Bild des Tieres und der Auswahl einer bestimmten Tierklasse durch Vorwissen des Betrachters ausgelöst wird. Der Kode existiert also bereits im Gehirn des Betrachters und gehört zu einer Menge von Kodes, die das Allgemeingut einer gesellschaftlich geprägten Denkweise darstellen.

5.5 Die spezielle Natur von Geo-Informationen

Geoinformationen oder raumbezogene Informationen sind Forschungsgegenstand einer jungen Wissenschaft, der *Geographical Information Science*. Der Grund hierfür ist, dass Geoinformation einen so eigenen Charakter besitzt, dass die mit ihr assoziierten Probleme einer gesonderten Behandlung bedürfen. Einen Eindruck der mit der Verarbeitung von Geoinformationen einhergehenden Probleme hat das erst vor kurzem gelöste Problem der Abbildung in herkömmlichen, relationalen Datenbanksystemen gegeben. Geoinformationen können auch als raumbezogene Informationen bezeichnet werden, was durch die „Geo“-Vorsilbe ausgedrückt wird. Somit lässt sich folgende Definition formulieren:

Geoinformationen zeichnen sich dadurch aus, dass eine der an der Information beteiligten Komponenten einen räumlichen Bezug bezeichnet.

Geoinformationen zeichnen sich daher gegenüber allgemeinen Informationen durch eine spezielle Komponente aus (Abb. 14.). Diese Komponente ist der Bezug durch eine Verortung auf der Erdoberfläche, die entweder explizit durch Angabe von Koordinaten und einem Raumbezugssystem ausgeprägt sein kann oder implizit durch eine topologische Beziehung, einen Ortsnamen oder ähnliches.

Geoinformationen können somit herangezogen werden, um Entscheidungen, die mit raumbasierten Phänomenen und Strukturen in Beziehung stehen, zu

verbessern. Es wird inzwischen allgemein davon ausgegangen, dass ein Großteil aller Planungsentscheidungen durch räumliche Information verbessert werden kann.

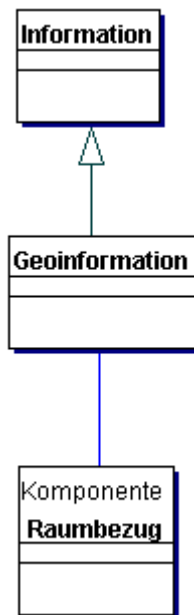


Abbildung 14. Geoinformation

Nach BERTIN (1974) ist das Besondere an Karten, dass die Beziehungen in der Ebene zwischen allen Teilen der gleichen Komponente zum Ausdruck kommen können. Diese Erkenntnis ist auf alle Arten von Geoinformation übertragbar und gibt uns so einen Hinweis auf den besonderen Charakter von Geoinformationen. Durch die Georeferenzierung, das heißt die Herstellung eines räumlichen Bezuges, ist es möglich, Beziehungen zu jedem anderen raumbezogenen Phänomen zu erstellen. Diese Beziehung lässt sich metrisch topologisch oder mengenbasiert darstellen (WORBOYS 1995).

Beispiel: Zur Verdeutlichung soll eine Karte der Verteilung des Ozonmaximums von 1995 herangezogen werden (Abb. 15.). Die Darstellung der Ozonmaxima in ihrer räumlichen Verteilung erlaubt die Erstellung vielfältiger Beziehungen. Innerhalb der Karte wird die Invariante Ozonmaximum mit Hilfe der geografischen Komponente und der nicht-geografischen Komponente Messwert des Ozonmaximums dargestellt. Die Kartendarstellung erlaubt es, Beziehungen zwischen allen Messwerten des Ozonmaximums herzustellen. Zusätzlich wird durch die Georeferenzierung

die Möglichkeit gegeben, Beziehungen zu beliebigen anderen verräumlichten Phänomenen aufzubauen, beispielsweise zu Zahlen der Verkehrsdichte, anderer Schadstoffemissionen oder der Häufigkeit von Lungenerkrankungen.

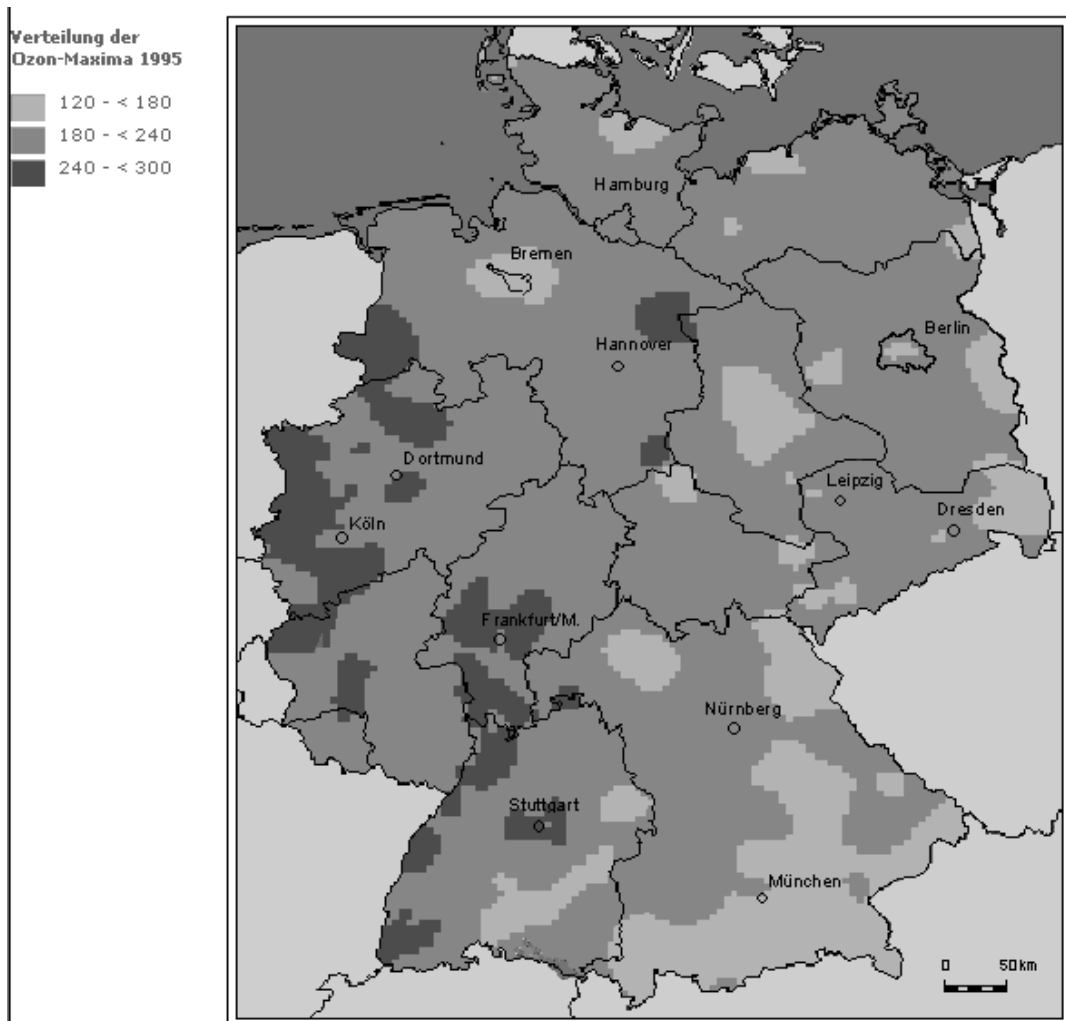


Abbildung 15. Verteilung des Ozonmaximums in der BRD 1995 (Quelle: Senatsverwaltung für Stadtentwicklung, Umweltschutz und Technologie 1996²⁸)

Der geografische Raum, der den Bezugsrahmen für Geoinformationen definiert, kann auf zwei unterschiedliche Weisen betrachtet werden, die sich auf der Datenebene widerspiegeln (COUCLELIS 1992). Zum Einen kann Raum als eine Sammlung von diskreten Objekten verstanden werden, die unterscheidbar sind und definierte Eigenschaften besitzen. Diese „herkömmliche“ Sichtweise des Raumes wird bevorzugt für Phänomene bis zu einer Größe, die für den Einzelnen noch direkt visuell erfassbar sind, wie beispielsweise Tische, Bäume oder Häuser. Ab einer gewissen Größe neigen

²⁸ siehe <http://www.stadtentwicklung.berlin.de/umwelt/umweltatlas/eia306.htm>

Menschen zu einer eher feldbasierten Denkweise, das heißt Phänomene werden weniger als diskrete Objekte, sondern als Verteilungen bestimmter Eigenschaften angesehen. Landschaften sind Beispiele für Phänomene dieser Größenordnung.

Diese Neigung zu einer speziellen Sichtweise geografischer Phänomene spiegelt sich bis zu einem gewissen Grade in den beiden Modellen geografischer Information, dem Feld- und dem Objektmodell (WORBOYS 1995) wider. Das Feldmodell betrachtet räumliche Information als Sammlungen von räumlichen Verteilungen, die als mathematische Funktionen formalisiert sein können. Mittels einer solchen Funktion wird jedem Punkt im Raum ein Attributwert zugewiesen. Auf Implementationsebene entspricht dies Daten, die beispielsweise als Raster oder *Triangulated Irregular Network* (TIN) modelliert wurden. Im Rahmen des Objektmodells werden Geoinformationen als diskrete Objekte behandelt, die jedes für sich einen Raumbezug besitzen, was auf Implementationsebene in der Regel dem Vektormodell entspricht. Aufbauend auf diesen Erkenntnissen und der schon aufgestellten Definition für Daten ist es somit möglich, folgende Aussage über Geodaten aufzustellen:

Geodaten sind atomisierte Geoinformationen, also Daten mit räumlichem Bezug. Dieser wird durch eines der beiden Modelle räumlicher Information erstellt, also entweder feld- oder objektbasiert.

Die Speicherung des Raumbezuges von Geodaten erfolgt im Falle objektbasierter Modellierung oftmals getrennt von den Sachdaten, zumindest falls diese in einer Datenbank gespeichert sind. Dieses „georelationale Modell“ wird inzwischen mehr und mehr von einer integrierten Speicherung von Geometrie- und Sachdaten innerhalb einer Datenbank abgelöst. Implementierungen des Feldmodells, vor allem Rasterdaten, sehen dagegen oftmals keinerlei Trennung von Geometrie- und Sachdaten vor. Mittels einer Wertematrix und Angaben über Rasterweite und Raumbezugssystem wird der vollständige Datensatz innerhalb einer Datei gespeichert.

Aus digitalen Geodaten können mittels eines GIS oder Computerkartographiesystems digitale Karten erstellt werden. Karten, sowohl

digitale als auch analoge, stellen grafische Zeichen dar, deren Signifikat raumbezogene Phänomene sind. Der Kode einer Karte setzt sich aus den einzelnen Signaturen zusammen, die mittels einer Legende erläutert werden, und dem Maßstab, welcher die Beziehung zwischen Karteneinheiten und tatsächlichen Raumeinheiten vermittelt.

5.6 Exkurs: Arten von Raumbezügen

Im folgenden Abschnitt sollen die Möglichkeiten Raumbezüge zu definieren kurz erläutert werden. Die Angabe der Verortung von Phänomenen lässt sich auf zwei grundsätzlich verschiedene Weisen gestalten, die eingeschränkt aufeinander abbildbar sind. Die beiden Möglichkeiten können als direkte oder indirekte Angabe des Raumbezuges bezeichnet werden. Eine direkte Angabe erfolgt über ein Raumbezugssystem, das aus einem Koordinatensystem, einer Projektion und einem Datum besteht sowie einer Menge von Koordinaten. Durch die Möglichkeit, Koordinaten jedes beliebigen Raumbezugssystems durch eine mathematische Transformation in jedes beliebige andere umzuwandeln, ist die räumliche Beziehung eines Phänomens zu jedem anderen durch diese direkte Raumbezugsangabe definiert. Die indirekte Angabe eines Raumbezuges kann sehr unterschiedlich erfolgen. Möglichkeiten hierfür sind die Angabe eines Ortsnamens, einer Adresse oder der räumlichen Beziehung zu einem verorteten Phänomen, wobei diese räumliche Beziehung durch topologische Parameter angegeben werden kann. Systeme, die indirekte Raumbezüge in einer thesaurusähnlichen Struktur verwalten, werden Geothesauri genannt (RIEKERT & TREFFLER 2000).

Die Abbildung zwischen direkten und indirekten Raumbezügen ist nur eingeschränkt durchführbar, da indirekte Bezüge oftmals nicht eindeutig lokalisierbar sind. Dies gilt für alle Arten indirekter Raumbezüge. Eine topologische Beziehung „benachbart zu“ gibt zwar eine grundlegende Auskunft darüber, wo sich ein verortetes Objekt befindet, aber noch keine eindeutige Auskunft über dessen Erstreckung. Gleiches kann für naturräumliche Gliederungen gesagt werden, da diese meist keine exakt festgelegten Grenzen besitzen. Die Abbildung zwischen direkten und

indirekten Raumbezügen unter Berücksichtigung ihrer zeitlichen Komponente wird durch Raumbezugsdienste zur Verfügung gestellt (FITZKE 2001).

5.7 Umweltinformationen und Geoinformationen

Wie bereits festgestellt, sind Geoinformationen Informationen mit Raumbezug. Umweltinformationen sind Informationen über die Umwelt. Eine genauere und ausreichend exakte Definition gibt das deutsche Umweltinformationsgesetz in §3 (2):

„Informationen über die Umwelt sind alle in Schrift, Bild oder auf sonstigen Informationsträgern vorliegenden Daten über

- 1. den Zustand der Gewässer, der Luft, des Bodens, der Tier- und Pflanzenwelt und der natürlichen Lebensräume,**
- 2. Tätigkeiten, einschließlich solcher, von denen Belästigungen wie beispielsweise Lärm ausgehen, oder Maßnahmen, die diesen Zustand beeinträchtigen oder beeinträchtigen können und**
- 3. Tätigkeiten oder Maßnahmen zum Schutz dieser Umweltbereiche einschließlich verwaltungstechnischer Maßnahmen und Programme zum Umweltschutz.“**

Jegliche Art von Umweltinformation kann erst durch Beziehung zu ihrer örtlichen Lage sinnvoll genutzt werden, eine Umweltinformation ohne Raumbezug ist entweder sinnlos oder allgemein gültig. Zur Lösung von Umweltproblemen ist eine geografische Komponente der Umweltinformation unbedingt notwendig. Umweltinformationen können daher als Spezialisierung der Geoinformationen betrachtet werden (siehe Abb. 16.). Die Untersuchung von raumbezogener Metainformation ist somit direkt für die Verarbeitung von Umweltinformationen relevant und auf Umweltdaten übertragbar.

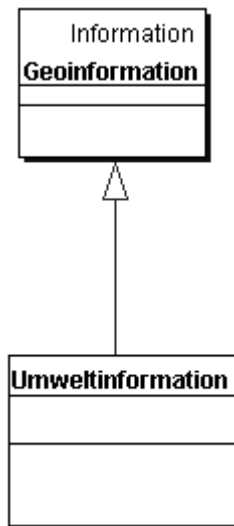


Abbildung 16. Umweltinformation

5.8 Metainformation und Metadaten

"Meta" bedeutet im ursprünglichen Wortsinn „zwischen“, „inmitten“, „nach“, „später“, wird aber auch im Sinne von „über“ und „danach“ verwendet (WAHRIG 1999). Metainformation könnte man somit als Information ansehen, die „danach“ erzeugt wird, im Sinne von nach der Erzeugung der eigentlichen Information. Allgemein wird der Ausdruck Metainformation auf alle einen Datensatz oder einen Dienst beschreibenden Informationen angewandt.

Die nun folgenden Setzungen bauen auf den Erkenntnissen des letzten Kapitels auf. Im Gegensatz zu allen bislang formulierten Definitionen sind diese Aussagen nicht direkt aus der Literatur entnommen oder abgeleitet. Die in Kapitel 5 identifizierten Funktionen von Metainformation dienen als Grundlage der weiteren Überlegungen zusammen mit den Konzepten der Semiotik. Obwohl die folgenden Überlegungen größtenteils auch auf nicht-raumbezogene Metainformationen anwendbar sind, ist das Hauptaugenmerk auf die raumbezogene Metainformationsverarbeitung gelegt. Es wird der Versuch unternommen, die Funktionen von Metainformation zu übersetzen in Konzepte und Funktionen von Zeichensystemen, um zu einem stabileren Fundament der Definition von Metainformation zu gelangen. Die Anwendung der semiotischen Begriffe hilft darüber hinaus, eine funktionale Gliederung von Metainformation aufzubauen.

Ein Datensatz wird dadurch erst zum Metadatensatz, indem er zur Beurteilung eines Datensatzes für einen wie auch immer gearteten Zweck benutzt wird (BREHERTON 1994). Würde er direkt zur Bearbeitung einer Fragestellung benutzt, wäre er Teil des Datensatzes. Wird ein Datensatz zur Beurteilung einer anderen Information oder eines anderen Datensatzes herangezogen, wird er zum Metadatensatz. Eine Information wird also zu einer Metainformation, indem sie losgelöst von der Information, mit der sie assoziiert ist, benutzt wird.

Die Fragestellung, beziehungsweise der Benutzungskontext, definiert somit die Natur eines Datensatzes (GÜNTHER 1998, KRASEMANN 1996). Folgende Fragestellungen können bei Nutzung eines Datensatzes unterschieden werden:

- Darstellung eines Phänomens oder der Beziehung zwischen Phänomenen
- Auskunft über ein Phänomen oder eine Beziehung zwischen Phänomenen
- Analyse eines Phänomens oder der Beziehung zwischen Phänomenen.

Diese entsprechen den raumbezogenen Basisoperationen (AUGSTEIN & GREVE 1994). Wird eine Information für eine dieser Fragestellungen benutzt, handelt es sich um eine Nutzung der Information als herkömmliche (= nicht-Meta) Information. Dies erlaubt im Umkehrschluss eine erste Definition von Metainformation: Metainformation ist die Information, die nicht für die Zwecke der gerade beschriebenen raumbezogenen Basisoperationen benutzt wird.

Eine der drei im letzten Kapitel identifizierten Aufgaben von Metainformation ist die Beschreibung des Entstehungskontextes, um ihn der Wiederverwendung durch Andere zugänglich zu machen. Diese Wiederverwendung entspricht der Absicht, den Datensatz für Zwecke der drei Basisoperationen anzuwenden. Wird eine Information benötigt, um einen schon identifizierten Datensatz für eine dieser Fragestellungen zu nutzen, wird sie als Kontextinformation benutzt. Sie stellt nach der semiotischen Begrifflichkeit den Kode dar, der notwendig ist, um dem Betrachter (im Falle einer analogen Karte) oder dem Systembenutzer (bei einer digitalen Karte oder einem Geografischen Informationssystem) erlaubt, die Information hinter der

Darstellung zu entschlüsseln. Beispiele für raumbezogene Kontextinformationen sind der Maßstab oder das Raumbezugssystem eines Geodatensatzes. Nicht spezifisch raumbezogene Metainformationen beschreiben beispielsweise den Mess- oder Herstellungsvorgang eines Datensatzes oder den Grund seiner Erstellung. Zusammenfassend kann definiert werden:

Kontextinformationen sind notwendig, um Datensätze für eine bestimmte Fragestellung nutzbar zu machen. Im Falle raumbezogener Informationen entsprechen die raumbezogenen Basisoperationen der Nutzarmachung des Datensatzes. Kontextinformationen bilden eine Spezialisierung von Metainformationen. Aus semiotischer Sicht stellen sie den Kode eines Geodatensatzes dar.

Zusätzlich existieren Metainformationen, die zur Benutzung eines Geodatensatzes notwendig sind, aber nicht für den menschlichen Benutzer gedacht und notwendig sind. Dies sind Informationen, die aus der informationstechnischen Speicherstruktur der Datensätze abgeleitet werden, wie beispielsweise Datentypen oder Dateitypen. Informationen dieser Art sollen als Strukturinformationen bezeichnet werden.

Strukturinformationen stellen den Kode dar, den eine Maschine benötigt, um einen Datensatz zu interpretieren. Informationen dieser Art sind implizit vorhanden als Struktur eines Datensatzes und können daraus abgeleitet werden.

Abb. 17. veranschaulicht den Kommunikationsprozess, dessen Ziel die Vermittlung von Geoinformationen ist, unter Berücksichtigung der Rolle, die Strukturdaten und Kontextdaten hierbei spielen.

Klassischerweise besteht ein Kommunikationsprozess aus den Komponenten Quelle – Sender – Kanal – Botschaft – Empfänger (ECO 1977). Die Quelle der Information beim digitalen Karten-Kommunikationsprozess ist ein raumbezogenes Phänomen. Durch Messung oder Abbildung des Phänomens werden Geodaten erzeugt. Dieser Abbildungsprozess kann beispielsweise durch Messung im Gelände erfolgen, mittels eines GPS oder eines Fernerkundungssensors. Falls Geodaten aus analogen Medien wie einer Karte

durch Digitalisierung abgeleitet werden, ist die Entstehung der analogen Karten vor der Erzeugung der digitalen Geodaten zwischengeschaltet. Kontextinformationen, die diesen Abbildungsprozess beschreiben, existieren bereits zu diesem Zeitpunkt. Falls bei der Erstellung digitaler Geodaten spezielle Erfassungsvorschriften benutzt werden, sind diese als standardisierte Kontextinformationen zu betrachten. Solche Erfassungs- oder Messvorschriften beschreiben die Vorgehensweise bei der Erzeugung digitaler Daten und deren Abbildung in einem Computermodell. Gleichgültig ob mittels Erfassungsvorschriften standardisiert oder für den speziellen Fall dokumentiert, entstehen somit Kontextdaten mehr oder weniger zeitgleich zu den Originaldaten. Zudem entstehen durch die maschinelle Speicherung und Kodierung der Geoinformationen in Geodaten die Strukturdaten.

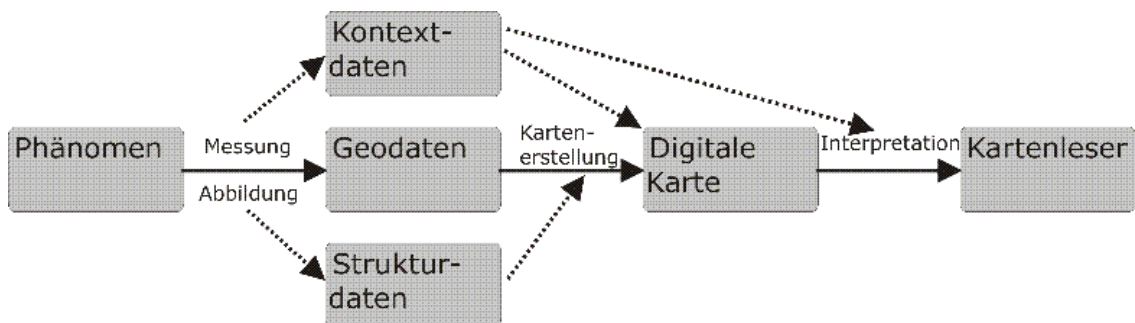


Abbildung 17. Kommunikationsweg der Geoinformation

Die Verarbeitung der Geodaten in einer Karte kann nun von den gleichen Personen durchgeführt werden, welche die Daten erhoben haben. In diesem Falle besitzen diese Personen bereits die Kontextdaten, die zur Erläuterung der Karte notwendig sind und hier in der Legende, dem Maßstab und zusätzlichen Erläuterungen ihren Platz finden. Strukturdaten werden vom benutzten digitalen Verarbeitungssystem benötigt, um die Geodaten verarbeiten zu können. Falls der Kartenerstellungsprozess nicht von den Datenherstellern durchgeführt wird, ist es notwendig, nicht nur die Geodaten selbst, sondern auch Strukturdaten und Kontextdaten an den Kartografen weiterzugeben, da sonst keine sinnvolle Weiterverarbeitung möglich ist. Strukturdaten müssen hierbei meist nicht explizit behandelt werden, da diese beispielsweise durch Dateieindungen weitergegeben werden. Die Weitergabe von Kontextdaten ist allerdings zwingend notwendig, um die Kartenerstellung zu ermöglichen. Falls

Kartenhersteller und -leser identisch sind, müssen die Kontextdaten nicht weiterverarbeitet werden. Dies ist insbesondere der Fall, wenn der Kartenleser – das heißt der Empfänger der Geoinformation - eine Datenanalyse durchführt. Die hierbei erzeugten Kartendarstellungen werden oftmals nicht explizit mit einer Legende ausgestattet. Kontextinformationen können somit auch direkt in den Interpretationsprozess einfließen. Darüber hinaus besteht die Möglichkeit, dass Kontextdaten nicht in die Legende mit eingehen und zusätzlich bei der inhaltlichen Interpretation herangezogen werden.

Darstellung und Analyse dienen der Erzeugung von Informationen, die zur Entscheidungsfindung benutzt werden können. Das Auffinden einer Information ist daher eine Fragestellung, die zeitlich früher angesiedelt ist. Diese vorgeordnete Fragestellung dient dem Finden von geeigneten Informationen für die eigentliche Frage. Werden Informationen in einem solchen Fragestellungskontext benutzt, werden sie als Kataloginformationen bezeichnet. Kontextinformationen können ebenfalls zur Navigation in Geodatenbeständen benutzt werden und werden somit zu Kataloginformationen.

Kataloginformation dient der Entscheidung über die Relevanz eines Datensatzes für eine spezifische Fragestellung. Jegliche Daten, die in diesem Sinne benutzt werden, werden als Katalogdaten bezeichnet, unabhängig davon, ob sie spezifisch hierfür erzeugt wurden oder Teil anderer Daten sind. Kataloginformationen sind eine Untergruppe von Metainformationen. Sie entsprechen einem Zeichen, das anstelle des eigentlichen Datensatzes betrachtet wird.

Die Definition der drei Typen von Metainformationen ist funktional aufgebaut. Es besteht folglich die Notwendigkeit, abhängig von ihrem Einsatz zu entscheiden, ob eine Information eine Struktur-, Kontext-, oder Kataloginformation und somit eine Metainformation ist. Wird sie für keine der drei genannten Funktionen genutzt, ist sie keine Metainformation. Zur Vereinfachung des Sprachgebrauches werden Nicht-Metainformationen im Weiteren als Basisinformationen bezeichnet.

Da die Unterscheidung der Arten von Metainformationen funktional getroffen wird, kann eine Information abhängig vom Anwendungskontext zu mehreren dieser Typen gehören, die Ergebnismengen der Unterscheidung sind demnach nicht disjunkt. Sowohl Struktur- als auch Kontextinformationen können zur Navigationsunterstützung in Geodatenansammlungen benutzt werden und werden somit zu Kataloginformationen. Darüber hinaus gibt es Metainformationen, die speziell zur Katalogisierung in solchen Datenbeständen erzeugt werden, um die Entscheidung über die Relevanz eines Geodatensatzes für eine Fragestellung zu unterstützen. Ein Beispiel für diese Art von Katalogdaten sind Zugriffsrechteinformationen oder Preisinformationen, das heißt Daten, die sich ausdrücklich mit der Abgabe von Datensätzen befassen. Diese Art von Informationen werden oftmals speziell für einen Datenkatalog erzeugt und sind insofern reine Kataloginformationen. Oftmals sind Kataloginformationen aber auch aus den Basisdaten abgeleitete Informationen (FLEWELLING & EGENHOFER 1999), wie beispielsweise Teile der katalogisierten Datensätze. Dem Recherchierenden werden in diesem Fall Teile der Basisdaten zur Verfügung gestellt, um die Entscheidung über die Relevanz eines Datensatzes zu unterstützen. Alle Arten von Vorschaubildern wie beispielsweise *Thumbnails* oder die bereits erwähnten *Footprints* des GIS-Pool Ansatzes, stellen ebenfalls Ansätze dar, bei denen Teile der Basisdaten zur Erstellung von Kataloginformationen genutzt werden. Ein Metadatensatz besteht folglich aus Strukturinformationen, Kontextinformationen, Teilen des beschriebenen Datensatzes und reinen Kataloginformationen. Er ist somit ein aus dem Code des Zeichensystems und zusätzlichen, unter Umständen aus dem Datensatz abgeleiteten, Informationen bestehendes Zeichen, das auf andere Zeichen ("echte" Datensätze) verweist. Die teilautomatisierte Erfassung von Metainformation (FRETTER 2000) entspricht somit der Umwandlung von Teilen der Basisdaten, Kontextdaten und Strukturdaten in Katalogdaten.

Eine spezielle Art von Metadaten sind Daten, die Dienste beschreiben. Dienste beschreibende Metadaten sind zum Teil nicht dafür gedacht, von Menschen gelesen zu werden. Diese technischen Metadaten erlauben es Programmen, welche die zur Verfügung gestellten Dienste nutzen wollen, diese zu identifizieren und anzusprechen. Dienste beschreibende Metadaten

ermöglichen es verteilten Systemen zu identifizieren, welche Funktionalitäten durch einen Dienst zur Verfügung gestellt werden. Sie können genau wie datenbeschreibende Metadaten in Katalogen verwaltet werden, wobei eine Unterscheidung notwendig erscheint zwischen einerseits an spezielle Datensätze gebundenen und andererseits freien Dienste-Metadaten.

Dienste beschreibende Metadaten sind das dynamische Äquivalent zu Daten beschreibenden Katalogdaten. Semiotisch gesehen ergibt sich kein Unterschied, da das beschriebene Zeichen kein Datensatz sondern ein Dienst ist. Dienste-Metadaten können an einen spezifischen Datensatz gebunden oder frei sein.

Der Unterschied zwischen einer sich auf einen Dienst und einen Datensatz beziehenden Metainformation liegt also in der Dynamik des Signifikats. Ein Dienst stellt ein informationstragendes Zeichen zur Verfügung, das noch nicht existiert, aber mittels des Dienstes zum Zeitpunkt der Anforderung erzeugt wird. Dieses Zeichen kann eine digitale Karte sein oder eine andere Art von Lageinformation, wie sie zum Beispiel für *Location based Services* benutzt wird.

Beispiel: Ein Benutzer eines internetbasierten Kartendienstes möchte eine Karte, auf der die aktuelle Ozonbelastung an seinem Wohnort dargestellt wird. Dieses Zeichen wird mit Hilfe eines Dienstes zur Verfügung gestellt, der aus Basisdaten (topografische Karte und Messdaten über die Ozonbelastung) erst zum Zeitpunkt der Anfrage die dementsprechende Information erzeugt. Um diesen Dienst zur Verfügung zu stellen, müssen die folgenden Metadaten bereitgestellt werden:

- verfügbare Themen
- unterstützte Raumbezugssysteme
- räumliche Abdeckung
- zeitliche Abdeckung und
- Zugriffsbeschreibung auf den Dienst.

Nur für den Fall, dass beide Datensätze den selben Zeitraum abdecken, im gleichen Raumbezugssystem und mit der gleichen Abdeckung existieren, kann die angeforderte digitale Karte erzeugt werden. Diese Überprüfung kann mit

Hilfe der bereitgestellten Metainformationen vorgenommen werden und gegebenenfalls kann, falls der Kartendienst dies unterstützt, ein Datensatz in das Raumbezugssystem des anderen überführt werden.

Bislang wurden Metainformationen beispielsweise von DENZER & GÜTLER (1994) im Kontext der Integration von Umweltdaten typisiert. Sie unterscheiden zwischen syntaktischer, semantischer, struktureller und navigatorischer Metainformation. Diese Typisierung basiert auf Schwierigkeiten bei der Datenintegration aufgrund von Heterogenität. Semantische Metainformation entspricht in dem hier vorgestellten Ansatz in etwa der Kontextinformation, die strukturelle dem gleichlautenden Begriff in dieser Arbeit. Sowohl semantische, syntaktische, strukturelle, als auch navigatorische Metainformation können Kataloginformation sein. An diese Typisierung angelehnt, aber mit leicht anderer Begrifflichkeit, unterscheiden GREVE & HÄUSLEIN (1994) zwischen semantischer, syntaktischer und pragmatischer Metainformation – zur Bildung von Kategorien werden Teilaspekte von Inhalten von Metainformationen benutzt. Die selben drei Kategorien wurden von MORRIS (1946, zitiert nach ECO 1977) als mögliche Betrachtungsweisen des Zeichens vorgeschlagen. Semantische Metainformation umfasst die inhaltliche Beschreibung von Datensätzen, syntaktische die formale und technische. Da Kontextinformationen die semantischen Äquivalenzen zwischen Signifikant und Signifikat erzeugen, gehören sie grundsätzlich zu den semantischen Metainformationen. Kataloginformationen umfassen syntaktische, semantische und pragmatische Metainformationen. Syntaktische Metainformation in den Ansätzen nach DENZER & GÜTLER (1994), als auch nach GREVE & HÄUSLEIN (1994) umfasst Informationen, die aus datenhaltungstechnischen Aspekten des zugrundeliegenden Datensatzes abgeleitet werden. Dies sind beispielsweise Daten- oder Dateitypen. Unterschiedlich werden in diesen beiden Ansätzen Informationen zu Verknüpfungen zwischen Daten benannt, im erstgenannten Ansatz als strukturelle Metainformation, im zweiten ebenfalls als syntaktische. Der *Data Dictionary*-Begriff, wie er im Umfeld von Datenbanken benutzt wird, katalogisiert diese Art von Daten, das heißt syntaktische und strukturelle Metainformation. Bei dem in dieser Arbeit vorgeschlagenen Ansatz entspricht

dies der Strukturinformation. Ein klassisches *Data Dictionary* entspricht demnach einem Katalog, der ausschließlich Strukturinformation enthält.

Die hier vorgeschlagene Typisierung ist, im Gegensatz zu den beiden gerade diskutierten, nutzungsorientiert. Grundlage bildet nicht ein bestimmtes Problem oder ein Aspekt der Daten, sondern der Grund ihrer Benutzung. Durch diese nutzungsorientierte oder funktionale Sichtweise wird eine grundlegendere Analyse des Problems erhofft. Sie erlaubt eine dynamische Definition von Metainformation und ihrer Spezialisierungen, wodurch eine grundlegende Entscheidung über die Natur eines Datensatzes bei seiner Entstehung unnötig wird. Somit ist die folgende allgemeine Definition möglich:

Als Metainformation wird jegliche Information bezeichnet, die entweder als Kode eines Datensatzes oder als Zeichen für einen Datensatz benutzt wird. Metainformation kann entweder Strukturinformation sein, die dem Kode entspricht, den eine Maschine zu Verarbeitung und Verwaltung von Daten benutzt. Sie kann aber auch Kontextinformation sein, welche dem Kode des Zeichens Digitale Karte entspricht, oder Kataloginformation, welche einen übergeordneten Zeichenprozess ausdrückt. Metainformation wird zeitlich vor der Benutzung von Daten benötigt.

Nachdem die Natur von Metainformationen hiermit ausreichend festgeschrieben ist, wird eine entsprechende Definition für Metadaten benötigt. Um der dynamischen Natur von Metainformationen zu entsprechen, wird folgendes vorgeschlagen:

Metadaten sind Daten, die ausdrücklich für die Ableitung von Metainformationen gespeichert wurden.

Hierbei ist anzumerken, dass Metainformationen nicht aus eigentlichen Metadaten abgeleitet werden müssen, sondern beispielsweise Ergebnis einer Anfrage an Basisdaten sein können. Die vorgeschlagene Definition drückt lediglich aus, dass Datensätze, wie sie beispielsweise für Geodatenkataloge im Sinne des OGC benutzt werden, um die Recherche nach Basisdaten zu unterstützen und von den Originaldaten getrennt gespeichert sind (siehe Abb. 7.), als Metadaten bezeichnet werden können.

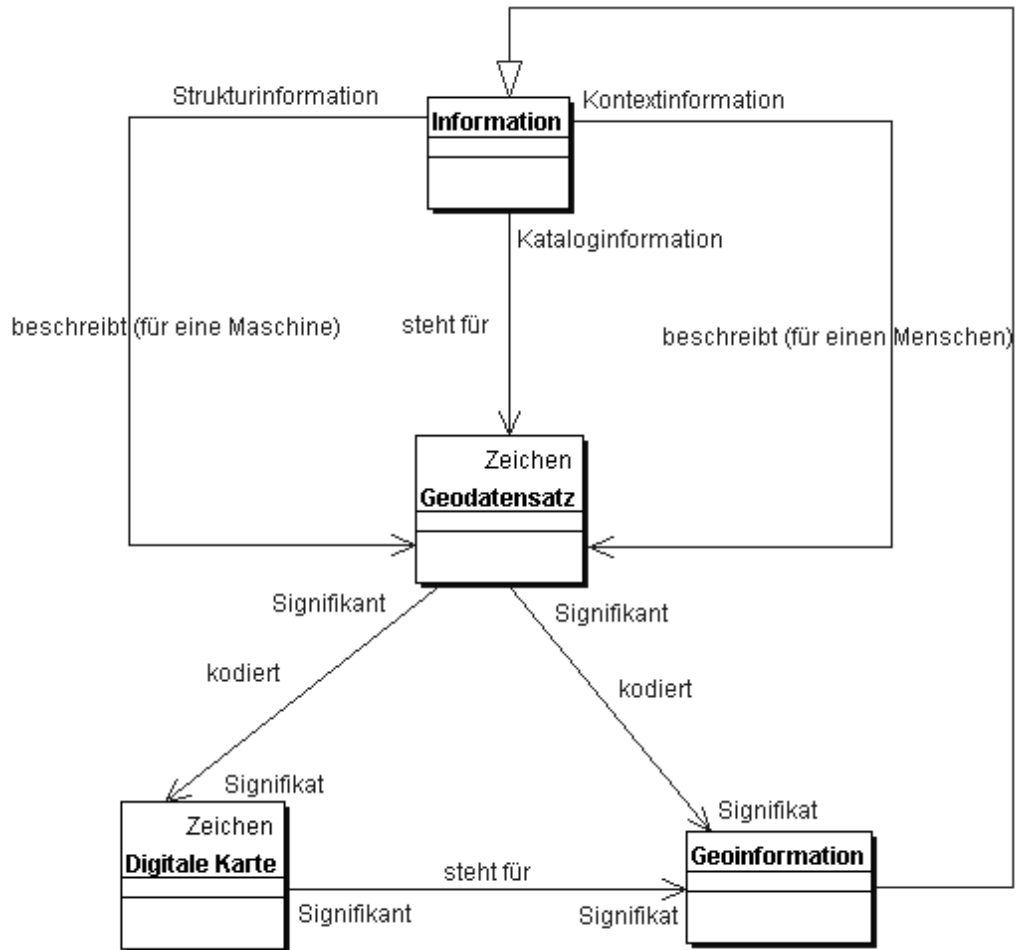


Abbildung 18. Rollen von Metainformation

Eine grafische Zusammenfassung der bisherigen Erkenntnisse diesen Abschnitts stellt Abb. 18. dar. Die Klasse Geoinformation ist wie bereits erwähnt eine Spezialisierung von Information. Im oberen Teil von Abb. 18. werden die verschiedenen Arten von Beziehungen dargestellt, die eine Information gegenüber einem Geodatenatz einnehmen kann und die sie zu einer Metainformation machen. Katalog-, Kontext- und Strukturinformation sind Rollen, welche durch die jeweiligen Funktionen der Information gekennzeichnet sind. Diese Rollen sind abhängig von der Art der Beziehung, welche die Information zum Geodatenatz besitzt. Im unteren Teil der Abbildung werden die Beziehungen der Klasse Geodatenatz zu Digitalen Karten und Geoinformation dargestellt, wobei in beiden Beziehungen Geodatenatz die Rolle des Signifikanten, also des Zeichens einnimmt. Beschreibt die Information den Geodatenatz für eine Maschine, das heißt um einem den Geodatenatz verarbeitenden Softwaresystem zu ermöglichen, aus

ihm eine Digitale Karte zu erstellen, dann ist die Information eine Strukturinformation. Dies entspricht der Beziehung zwischen Geodatenatz und Digitaler Karte, da eben diese Strukturinformation notwendig ist, um mittels eines GIS oder Computerkartografiesystems eine Karte zu erstellen. Aus semiotischer Sichtweise entspricht dies dem Kode, der die Entschlüsselung des Zeichens Geodatenatz ermöglicht.

Die gleiche Funktion, nur für einen Menschen, übernimmt die Kontextinformation. Sie erlaubt die Entschlüsselung der Semantik einer Digitalen Karte. Dies entspricht der Auflösung der Beziehung zwischen Geoinformation und Geodatenatz im unteren Teil der Abbildung. Ein Geodatenatz kodiert eine Geoinformation, diese kann auf prägnante Weise aber nur über den Umweg der Digitalen Karte entschlüsselt werden.

Als letzte noch nicht erläuterte Beziehung in Abb. 18. bleibt die mittlere der drei Beziehungen zwischen Information und Geodatenatz. Steht eine Information für einen Geodatenatz und erlaubt somit die Entscheidung über seine Relevanz für eine gegebene Fragestellung, dann ist sie eine Kataloginformation.

In Abb. 18. wurde darauf verzichtet, auf Dienste und deren Metadaten hinzuweisen. Sie ist auf diesen Fall in gleicher Weise anwendbar, nur dass der Prozess der Kartenerstellung unter Umständen durch mehrere Dienste ausgeführt wird.

Der Begriff Metainformation ist in Abb. 18. nicht mehr explizit aufgeführt. Metainformation ist ein Über- beziehungsweise Sammelbegriff für die drei Rollen Struktur-, Kontext- und Kataloginformation, wie in der Definition bereits ausgeführt.

In den bislang entwickelten Ausführungen besteht allerdings eine leichte Ungenauigkeit in Hinsicht auf den Umgang mit dem Begriff Kataloginformation, die nun noch aufgelöst werden soll. Die Aussage, dass Kataloginformation an Stelle eines Geodatenatzes steht, muss in der Hinsicht präzisiert werden, dass ein Katalogeintrag, welches ein Zeichen ist (siehe Abb. 19.), eine Kataloginformation kodiert.

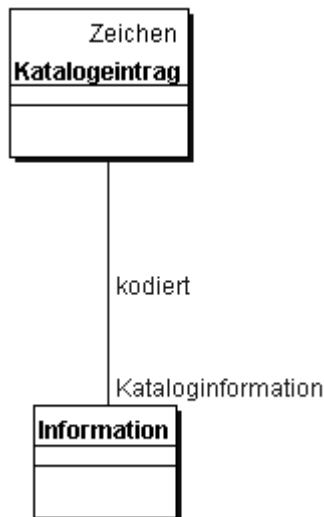


Abbildung 19. Katalogeintrag als Zeichen

Ein Katalogeintrag wird somit anstelle des Geodatenatzes, beziehungsweise einer mit Hilfe des Geodatenatzes erzeugten Digitalen Karte, betrachtet, um Informationen über den Geodatenatz zu erhalten. Die Klasse Katalogeintrag entspricht im Übrigen dem *CG_CatalogEntry* im Katalog Objektmodell der *Catalog Interface Implementation Specification* (Abb. 9.).

5.9 Raumbezogene Metainformation

Zur Nutzung von Geoinformationen sind spezielle Metainformationen notwendig, die über den Bedarf zur Verarbeitung anderer Informationen hinausgehen. Aufbauend auf den schon getroffenen allgemeinen Aussagen über Metainformation bezeichnet raumbezogene Metainformation somit Zeichen, deren Signifikat Geodatenätze, raumbezogene Dienste oder Karten sind, beziehungsweise den Kode eines Geodatenatzes. Kontextinformationen für Geoinformationen umfassen die Beschreibung des räumlichen Bezuges, das heißt jene Informationen, die in Karten mit Hilfe von Legende und Maßstab vermittelt werden. Geo-Metainformation ist eine Spezialisierung von allgemeiner Metainformation (siehe Abb.20.).

Im Falle von Kataloginformation stellt die Metainformation den Informationsgehalt eines Zeichens dar, welches stellvertretend für einen Geodatenatz steht. Dieses Zeichen wird als Katalogeintrag bezeichnet. Ein Zeichen für einen Geodatenatz wird sinnvoller Weise das charakteristische Merkmal eines raumbezogenen Datensatz abbilden, eben den Raumbezug.

Spezifisch raumbezogene Metainformationen sind somit Informationen über das „Wo“ eines Datensatzes, was im einfachsten Fall ein umgebendes Rechteck ist. Um dieses „Wo“ vollständig zu definieren, ist zusätzlich zu den Koordinaten das Raumbezugssystem anzugeben. Die räumliche Metainformation kann bedeutend komplexer strukturiert sein, als ein bloßes Koordinatenrechteck. Es kann hierfür zum Beispiel der Umriss eines Geodatensatzes, ein „Thumbnail“-Bild oder ein *Footprint* (GREVE, LESSING & MÜLLER 2001) benutzt werden. Allgemein kann jede Teilmenge oder Ableitung der Geometrie als Geo-Metainformation benutzt werden.

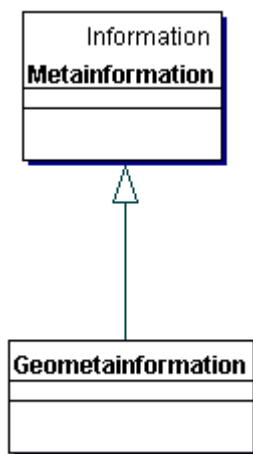


Abbildung 20. Geo-Metainformation

Alle diese Möglichkeiten zur Darstellung des Raumbezuges und somit der Erzeugung eines Zeichens, beziehen sich auf direkte Raumbezüge, das heißt Raumbezüge, die durch Koordinaten hergestellt werden. Im Falle einer indirekten Verortung eines Datensatzes/Dienstes kann diese unter Umständen direkt vom Basis-Datensatz in die Metainformation übernommen werden.

Bei der Verortung eines Datensatzes mittels eines Geothesaurs kann der selbe auch für eine Katalogrecherche benutzt werden. Hierbei stellt sich die Frage, welche der beteiligten Komponenten in diesem Falle die Metainformation darstellt. Wie schon ausgeführt, ist Metainformation entweder Information, die aus einem Zeichen, das für einen Geodatensatz steht, abgeleitet wird²⁹, oder der Kode, der das mittels eines Geodatensatzes erzeugte Zeichen³⁰ zum

²⁹ dies würde Kataloginformation entsprechen

³⁰ im Falle der raumbezogenen Informationsverarbeitung wäre dies eine Karte

Bezeichneten mittels semantischer Äquivalenzen in Bezug setzt³¹, oder maschinenverarbeitbare Information über den Aufbau des informationstragenden Datensatzes³². Folglich ist die Bezeichnung einer räumlichen Einheit das (Teil-) Zeichen, welches die Metainformation trägt. Aufbauend auf dieser Bezeichnung können Verknüpfungen zu weiteren räumlichen Einheiten mittels eines *Gazetteers* erstellt werden. Diese sind aber nicht mehr Teil des Meta-Zeichens Katalogeintrag.

Da raumbezogene Metainformation oder Geo-Metainformation eine Spezialisierung von allgemeiner Metainformation darstellt, ist folgende Definition naheliegend:

Raumbezogene Metainformation bezieht sich auf raumbezogene Datensätze oder Dienste. Sie definiert die Verortung eines Datensatzes oder einer Information.

Ein Katalogeintrag eines raumbezogenen Datensatzes besitzt somit ebenfalls Raumbezug und wird folglich zum *Feature* (siehe auch Abb. 9.). Ein solcher Katalogeintrag besteht aus textuellen und grafischen Beschreibungen des Datensatzes. Wie schon früher in dieser Arbeit ausgeführt haben räumliche Zeichensysteme Vorteile hinsichtlich der Informationsableitung gegenüber linearen Zeichensystemen wie der Sprache. Der Grund hierfür ist, dass sie Verbindungen zwischen drei Variablen, den beiden Dimensionen der Ebene und grafisch dargestellter Information, herzustellen ermöglichen. Raumbezogene Metainformation erlaubt also nicht nur die Erstellung zusätzlicher Informationen, sondern möglicherweise auch die effizientere Ableitung von Informationen. Hierzu ist es allerdings notwendig, eine kartografische Repräsentationsform zu benutzen, die Vorteile in der Prägnanz der Informationsübertragung gegenüber nichtgrafischen Darstellungsformen besitzt. Es kann somit die Empfehlung ausgesprochen werden, bei der Erstellung von Katalogeinträgen in Geodatenkatalogen den Raumbezug des Eintrages mittels einer kartografischen Form darzustellen, das heißt mittels einer statischen oder dynamischen Kartenkomponente. Die Prägnanz dieser

³¹ was Kontextinformation entsprechen würde

³² also Strukturinformation

Darstellungsform legt es nahe, nicht nur den reinen Raumbezug, sondern auch zusätzliche Informationen zum Datensatz, wie beispielsweise Aktualität der Daten oder Datendichte, auf diese Weise darzustellen.

5.10 Zusammenfassung

Die soeben durchgeführte Modellbildung für raumbezogene Metainformation baut auf den Konzepten der Semiotik auf. Die Begriffe Zeichen und Kode werden benutzt, um eine Unterscheidung zwischen Metainformation und Basisinformation festzulegen und die verschiedenen Typen von Metainformation zu kategorisieren. Grundlage dieser Kategorisierung stellen die Ergebnisse aus Kapitel 4 dar. Die Definition von Metainformation ist kontextabhängig, nur durch Art ihres Einsatzes kann über ihre Natur entschieden werden. Es existieren drei Arten von Metainformation, die durch ihre Rollen festgelegt werden:

- Strukturinformation stellt den maschinenlesbaren Kode eines Datensatzes dar, der benötigt wird um sie informationstechnisch zu verwalten und zu verarbeiten
- Kontextinformation ist der Kode, der benötigt wird, um die Semantik eines Geodatensatzes oder einer Geoinformation zu entschlüsseln
- Kataloginformation dient der Recherche nach Geoinformationsressourcen, seien dies Daten oder Dienste. Sie definiert Zeichen, die stellvertretend für die eigentliche Ressource betrachtet werden.

Wird eine Information auf eine der drei angegebenen Arten benutzt, handelt es sich um Metainformation.

Raumbezogene Metainformation stellt eine Spezialisierung von allgemeiner Metainformation dar, die eine raumbezogene Komponente umfasst. Raumbezogene Metainformation, beziehungsweise raumbezogene Kataloginformation, sollte durch eine kartografische Darstellungsform repräsentiert werden, um Vorteile in Hinsicht auf die Effizienz der Informationsübertragung zu gewährleisten.

6 Diskussion und Ausblick

Nachdem der theoriebildende Teil im vorangegangenen Kapitel abgeschlossen wurde, soll nun diskutiert werden, welchen Nutzen eine solche Theorie hat. Ferner wird die Bedeutung aktueller Entwicklungen in der raumbezogenen Metainformationsverarbeitung diskutiert und anhand eines konkreten Beispiels veranschaulicht. Abschließend werden Überlegungen angestellt, welche Rolle die Metainformationsverarbeitung in der Zukunft der Geoinformationsverarbeitung spielen wird.

6.1 Nutzen einer Theorie der raumbezogenen Metainformationsverarbeitung

Problemlösungen in jedem Bereich menschlichen Lebens erfordern Konsens von Seiten der Beteiligten. Dieser Konsens entsteht durch Harmonisierung der jeweiligen „Wirklichkeiten“ der Beteiligten, also der Einigung über Wahrnehmungsmodelle. Je komplexer, abstrakter und dynamischer eine Problemstellung ist, desto höher ist die Wahrscheinlichkeit, dass Personen konkurrierende Denkmodelle eines Phänomens besitzen. Wissenschaft an sich befasst sich in ihrem Kern mit der Entwicklung und Homogenisierung von solchen Denkmodellen („Theorien“), die sich durch die Bildung und Ablehnung von Hypothesen evolutionär entwickeln. Das Phänomen Metainformation ist durch seine abstrakte und dynamische Natur sehr komplex, darüber hinaus ein zentrales Thema der raumbezogenen Informationsverarbeitung und somit hochgradig praxisrelevant. Das Problem der Terminologie für Metainformation und verwandte Begriffe besteht vor allem darin, dass nicht per se zu entscheiden ist, ob eine Information, beziehungsweise die die Information speichernden Daten, als Metadaten zu bezeichnen sind. Nur unter Berücksichtigung ihres Benutzungskontextes, das heißt der Rolle, welche die Daten spielen, kann diese Entscheidung getroffen werden. Diese rollenbasierte Sichtweise bedarf einer Interpretation der Metainformationsverarbeitung beschreibenden Begriffe vor einem kommunikationstheoretischen Hintergrund. Diesen Hintergrund liefert die

Semiotik. Die Verwendung der Semiotik erweist sich zudem äußerst sinnvoll, da die Darstellung raumbezogener Phänomene mittels Karten ebenfalls auf den Aussagen der Semiotik fußt (HAKE, GRÜNREICH & MENG 2001). Eine auf diesen Konzepten aufbauende Terminologie ist Ergebnis dieser Arbeit.

Über die Entwicklung einer konsistenten Terminologie hinaus und auf ihr aufbauend ist ein weiteres Ergebnis dieser Arbeit, dass grafische Darstellungsformen von Metainformation die gleichen Vorteile aufweisen, wie dies allgemein grafische Zeichensysteme gegenüber linearen tun. Katalogeinträge, die metainformationstragende Zeichen sind, profitieren somit von grafischen Darstellungsformen. Diese haben gegenüber textuellen Darstellungen Vorteile in Hinsicht auf die Prägnanz der Darstellung und sind deshalb bevorzugt zu benutzen.

Die Diskussion über die Natur von Metadaten und Metainformationen und ihre Bedeutung wird in nahezu jedem größeren GI-Projekt früher oder später geführt. Eine einheitliche Terminologie und Denkweise wie sie in dieser Arbeit entwickelt wurde, kann die Kommunikation zwischen Planern, Entwicklern und Nutzern von Metainformationssystemen stark vereinfachen und zu Arbeitserleichterung führen. Wie gezeigt ist, um dem dynamischen Aspekt von Metainformation gerecht zu werden, eine funktionale Klassifikation notwendig, die über die Natur einer Information oder eines Datensatzes erst mit Hilfe des Anwendungskontextes entscheidet.

6.2 Gegenwärtige Entwicklungen im Bereich der raumbezogenen Metainformationsverarbeitung

6.2.1 Bedeutung von XML

Die Datenauszeichnungssprache XML hat sich inzwischen als das Austauschformat für Informationen nicht nur des Internets etabliert. Raumbezogene Metainformationen können in diesem Zusammenhang als eine Untergruppe dieser Informationen interpretiert werden, der Austausch mittels anderer Formate wird in Zukunft nicht mehr von Bedeutung sein. Das Problem der technischen Interoperabilität von Metainformationssystemen ist hierdurch stark vereinfacht worden. Die Verbreitung von XML äußert sich darin, dass

Metainformationssysteme XML entweder als Im- und Exportformat unterstützen, oder Metadaten direkt als XML entweder in Dateien oder speziellen XML-Datenbanken gespeichert werden. Die Erleichterung in Hinsicht auf den Datenaustausch und die syntaktische Interoperabilität ergibt sich direkt als Folge dieser Verbreitung, XML-Dokumente können mit Hilfe der XSLT-Technologie aufeinander abgebildet werden.

Im Moment zeichnet sich ab, dass *XML-Schema* die Benutzung von *Document Type Definitions* (DTD) verdrängen wird, XML aber die allgemein verbreitete Sprache zum Austausch von strukturierbaren Informationen der nächsten Jahre sein wird. Die relative Mühelosigkeit mit der sich XML-Daten verschiedener Schemata mittels XSLT ineinander überführen lassen, egal ob deren Struktur durch eine DTD oder ein XML-Schema definiert ist, ist ein starkes Argument für die Sinnhaftigkeit dieser Technologie.

Abgesehen von der weiten Verbreitung von XML, eignet sich diese Sprache speziell zur Speicherung von Metainformation. Der Grund hierfür liegt darin, dass XML, zumindest mittels XML-Schema, eine sich selbst beschreibende Sprache darstellt. Ein XML-Schema ist nichts anderes, als eine XML-Datei, welche die inhaltliche Struktur anderer XML-Dateien vorgibt. Dies steht im Gegensatz zur Verwaltung von Daten, wie sie beispielsweise in Datenbanken vorgenommen wird, bei der sich die Sprache zur Definition des Schemas und der eigentlichen Daten unterscheiden. XML macht somit alleine von seiner Struktur schon deutlich, dass die Unterscheidung zwischen Information und Metainformation nicht generell, sondern nur kontextabhängig zu treffen ist.

6.2.2 Standardisierung von Metadaten

Die Erreichung technischer Interoperabilität mittels XML stellt nur ein zu lösendes Problem auf dem Weg zu echter interoperabler Metainformationsverarbeitung dar. Die Beschreibung und Semantik der Metadatenelemente selbst differiert stark in den verschiedenen Ansätzen zur raumbezogenen Metainformationsverarbeitung. Aus diesem Grund entwickelten sich in den letzten Jahren diverse Metadaten-Standards, die eine einheitliche Verwendung von Metadaten zum Ziel haben. Die wichtigsten Initiativen dieser Art sind der CSDGM, ENV 12657, für den

deutschsprachigen Raum das Datenmodell des UDK und vor allem ISO 19115 (MÜLLER 2001a). Es wurde allerdings schon früh erkannt, dass eine zu starre Standardisierung nicht zum Ziel führt, da kein Metadatenstandard alle Anwendungen raumbezogener Metainformationsverarbeitung hinreichend genau abzudecken im Stande ist. Aus diesem Grunde wurde das Konzept optionaler Elemente und Profile eingeführt, die eine Anpassung eines Standards an die eigenen Bedürfnisse erlaubt.

Die Entwicklung von ISO 19115 ist als Meilenstein der raumbezogenen Metainformationsverarbeitung zu betrachten. Nahezu alle wichtigen internationalen Standards gingen in die Entwicklung mit ein und es zeichnet sich eine breite Akzeptanz des Standards ab, welche nicht zuletzt auf die Unterstützung durch das OGC zurückzuführen ist. Da die meisten Organisationen, die selbst an der Entwicklung eines Metadatenstandards arbeiteten, die Übernahme von ISO 19115 angekündigt haben, erscheint es wahrscheinlich, dass mittelfristig keine anderen Standards mehr unterstützt werden. Die Umsetzung durch die führenden Hersteller von GI-Software scheint ebenfalls gegeben. Als nur ein Beispiel sei angemerkt, dass ESRI, immer noch der Marktführer im Bereich von GIS-Software, die Umstellung von CSDGM auf ISO 19115 in den zukünftigen ArcGIS-Produkten angekündigt hat (ESRI 2001). Aktuell stattfindende Metainformationsinitiativen wie GeoMIS.Bund setzen von Anfang an auf diesen Standard.

Das semantische Problem der raumbezogenen Metainformationsverarbeitung wäre somit größtenteils gelöst. Es bleibt natürlich eine gewisse Unschärfe bei der Benutzung des ISO-Standards erhalten, da immer noch eine unterschiedliche semantische Interpretation der einzelnen Metadatenelemente möglich ist.

6.2.3 Interoperable Geodatenkataloge durch Unterstützung von OGC – Spezifikationen

Um echte Interoperabilität im Bereich der Geo-Metainformationsverarbeitung zu erlangen, genügt die Standardisierung durch XML und einen internationalen Standard allerdings noch nicht. Was hierzu noch nötig ist, ist die Unterstützung wohldefinierter Schnittstellen und Protokolle durch

Geodatenkataloge. Die *Catalog Interface Implementation Specification* des OGC (OGC 1999e) stellt hierfür die Grundlage dar. Die Umsetzung dieser Spezifikation wurde vermutlich durch die Unterstützung von CORBA, COM und Z39.50 eher behindert, erst das neue Web-Profil (REICH & PANAGIOTIS 2001) wird zu ihrem Durchbruch führen. Der Grund hierfür ist, dass sowohl CORBA- als auch COM-Verbindungen nur in einer sehr kleinen und wohldefinierten Netzumgebung sinnvoll sind und die Unterstützung des Z39.50 Protokolls sehr komplex ist. Web-Services dagegen sind allgemein zugänglich und auf vergleichsweise einfache Art zu implementieren. Das im nächsten Abschnitt beschriebene System basiert auf dem gerade erwähnten Web-Profil der *Catalog Interface Implementation Specification*.

6.3 Konzeption und Entwicklung eines Geodatenkatalogs am Beispiel des Metainformationssystems von Luxemburg

6.3.1 Allgemeines

Das Metainformationssystem für Geodaten des Großherzogtums Luxemburg (MISLux) dient als Beispiel für den Aufbau eines Geodatenkataloges, der sich aktueller Konzepte und Technologien bedient. Das Ziel von MISLux ist die Verfügbarmachung von Geodaten der öffentlichen Verwaltung für Nutzer in- und außerhalb dieser Verwaltung. Die Planung von MISLux hat in der Mitte des Jahres 2001 begonnen, Anfang 2002 wurde ein Prototyp fertiggestellt, welcher zusammen mit einem detailliert ausgearbeiteten Feinkonzept die Grundlage der weiteren Entwicklung darstellt.

MISLux ist der Arbeitsname dieses Metainformationssystems für Geodaten. Es basiert auf einem Konzept, welches die Integration von Geodaten aus den verschiedenen Verwaltungsabteilungen in einen Internet-basierten Geodatenkatalog vorsieht. Die Ziele von MISLux sind Datendokumentation und Verfügbarmachung von Daten innerhalb der Verwaltung Luxemburgs und darüber hinaus. MISLux stellt ein dezentrales Informationssystem dar, auf welches mittels eines zentralen Katalog-Servers (dem „Portal“) zugegriffen werden kann.

Durch die Nutzung internationaler Standards (*Open GIS Consortium, International Standards Organisation*) soll die Offenheit und Erweiterbarkeit des Systems gewährleistet werden. MISLux soll sowohl die Recherche auf Metadatenenebene erlauben, als auch den direkten Zugang zu digitalen Geodaten³³. Da die Produkte der Firma ESRI in den Verwaltungen von Luxemburg eine große Verbreitung gefunden haben, müssen diese eine besondere Berücksichtigung erhalten.

Die Entwicklung von MISLux wird in drei Phasen durchgeführt. In der ersten, inzwischen abgeschlossenen Phase, wurde ein Funktionsprototyp erstellt, der die Grundlage der weiteren Implementation darstellt. Ein eigens für die Prototypphase aufgesetztes beispielhaftes Fachinformationssystem dient als Datenquelle des Systems. In der zweiten Phase der MISLux-Entwicklung sollen die einzelnen Fachinformationssysteme erschlossen und die Funktionalität für den Einsatz innerhalb der Verwaltung Luxemburgs abgeschlossen werden. Zumindest die, durch die Mitglieder der den Auftraggeber darstellenden Koordinationsgruppe abgedeckten, sechs Fachbereiche sollen hierbei berücksichtigt werden. In der dritten Phase wird MISLux um eine E-commerce Komponente erweitert, die eine Abgabe von Daten an Dritte erst nach Zahlung einer Gebühr erlaubt. Außerdem werden weitere, noch zu identifizierende Abgabeformate unterstützt (beispielsweise DXF³⁴, GML³⁵).

6.3.2 Funktionsumfang

Der Funktionsumfang von MISLux lässt sich in Anlehnung an die *Catalog Service Implementation Specification* (OGC 1999e) in drei Kategorien unterteilen (Abb. 8.):

- Recherche (*Discovery Service*)
- Datenzugriff (*Access Service*)
- (Meta-)Datenmanagement (*Management Service*)

³³ Die Möglichkeit von Anbieterseite nur Metadaten einzustellen bleibt allerdings weiterhin erhalten.

³⁴ Data Exchange Format

³⁵ Geography Markup Language, eine XML-Sprache zur Beschreibung von Geodaten (OGC 2001c)

Die Recherche nach Geodatenätzen basiert auf freien Suchbegriffen und Auswahllisten von Suchtermen. Es werden sowohl fachliche, als auch räumliche und zeitliche Suchkriterien benutzt. Die räumliche Recherche wird durch eine interaktive Karte und Auswahllisten der Verwaltungseinheiten des Großherzogtums Luxemburg unterstützt. Die textbasierte Recherche erfolgt mittels Freitextsuche und Auswahllisten von Ordnungskategorien. Zur Unterstützung der Freitextsuche ist der Einsatz eines Fachthesaurus vorgesehen, die Wahl fiel hierbei auf den *General Multilingual Environmental Thesaurus* (GEMET), da dieser als einer der wenigen existierenden Thesauri die benötigte Mehrsprachigkeit aufweist (LETHEN 1999). Der eingesetzte Thesaurus hat im Rahmen der Recherche zwei Funktionen zu erfüllen. Zum ersten wird mit seiner Hilfe eine Übersetzung der Suchbegriffe in die anderen, innerhalb der Metadaten benutzten, Sprachen vorgenommen. Zum anderen kann er benutzt werden, um dem Anfragenden entsprechende Suchvokabeln vorzuschlagen für den Fall, dass dieser nur ungenaue Vorstellungen zur Einschränkung seiner Suchanfrage besitzt. Die Recherchekomponente ist über Internet öffentlich zugänglich und kann von jedermann eingesehen werden. Resultat einer Recherche ist eine ISO 19115-konforme Metadatenmenge zur Identifizierung relevanter Datensätze.

Im Falle der Identifikation eines für den Nutzer interessanten Datensatzes besteht die Möglichkeit, diesen anzufordern. Dies geschieht mittels Angabe eines Datensatzidentifikators, einer räumlichen Recherchebedingung und eines Abgabeformates. Das Resultat einer solchen Anfrage wird zum http-Download mittels einer URL³⁶ zur Verfügung gestellt. Der Anfragende muss sich vor dem Download mittels seines Login und Passwortes beim System identifizieren.

Das Metadatenmanagement umfasst die Möglichkeit, neue Datensätze beim MISLux-System anzumelden. Dies erfolgt mittels einer XML-Datei, die wahlweise mittels E-Mail an den MISLux-Server gesendet oder auf dem Server des Fachinformatiossystems zum automatischen Upload zum MISLux-System zur Verfügung gestellt wird. Das MISLux-Portal übernimmt die

³⁶ Uniform Resource Locator

Aufgabe, die Metadatensätze der verschiedenen Fachinformationssysteme in eine einheitliche und recherchierbare Datenmenge einzugliedern.

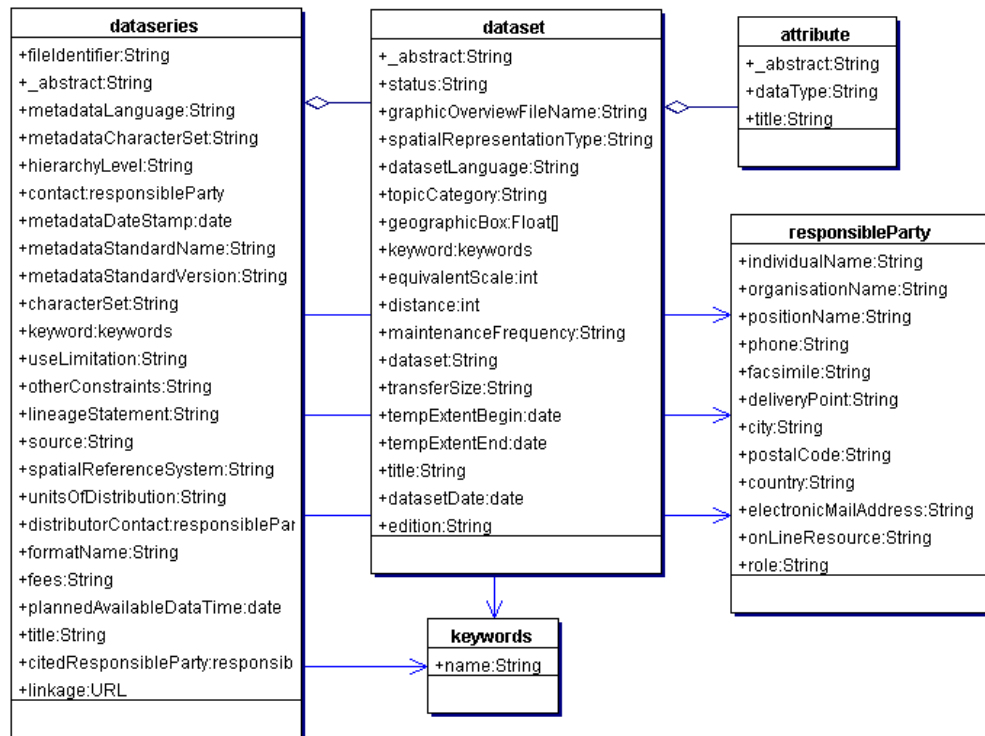


Abbildung 21. Metadatenmodell von MISLux in UML Notation

6.3.3 Datenmodell

Das Metadatenmodell von MISLux basiert auf dem internationalen Standard ISO 19115 *Geographic Information – Metadata*. Zusammen mit den zukünftigen Nutzern des Projektes wurden die erforderlichen Metadatenelemente identifiziert und in ein Datenmodell umgesetzt, das Abb. 21. entnommen werden kann. Es ist anzumerken, dass dieses Diagramm in der objektorientierten UML-Notation dargestellt ist, wobei die Datentypen vereinfacht sind. Es stellt sowohl die Grundlage des entwickelten XML-Schemas, als auch der Umsetzung in einer relationalen Datenbank dar. Im Falle der relationalen Umsetzung ist es notwendig, abhängig von der Kardinalität der Beziehungen zwischen den einzelnen Relationen, einige zusätzliche Relationen und Schlüsselattribute zu definieren.

Das Metadatenmodell unterteilt zwischen zwei Hierarchiestufen, es existieren Metadaten für einzelne Datensätze³⁷ als auch für übergeordnete Datenserien³⁸. Die Metadaten dieser übergeordneten Einheit sind für alle zugehörigen Datensätze gültig. Darüber hinaus existieren Adressinformationen³⁹, die in unterschiedlichen Rollen einer Datenserie zugeordnet werden können, eine Klasse zur Beschreibung von Attributen der Datensätze⁴⁰ und der Schlüsselworte⁴¹.

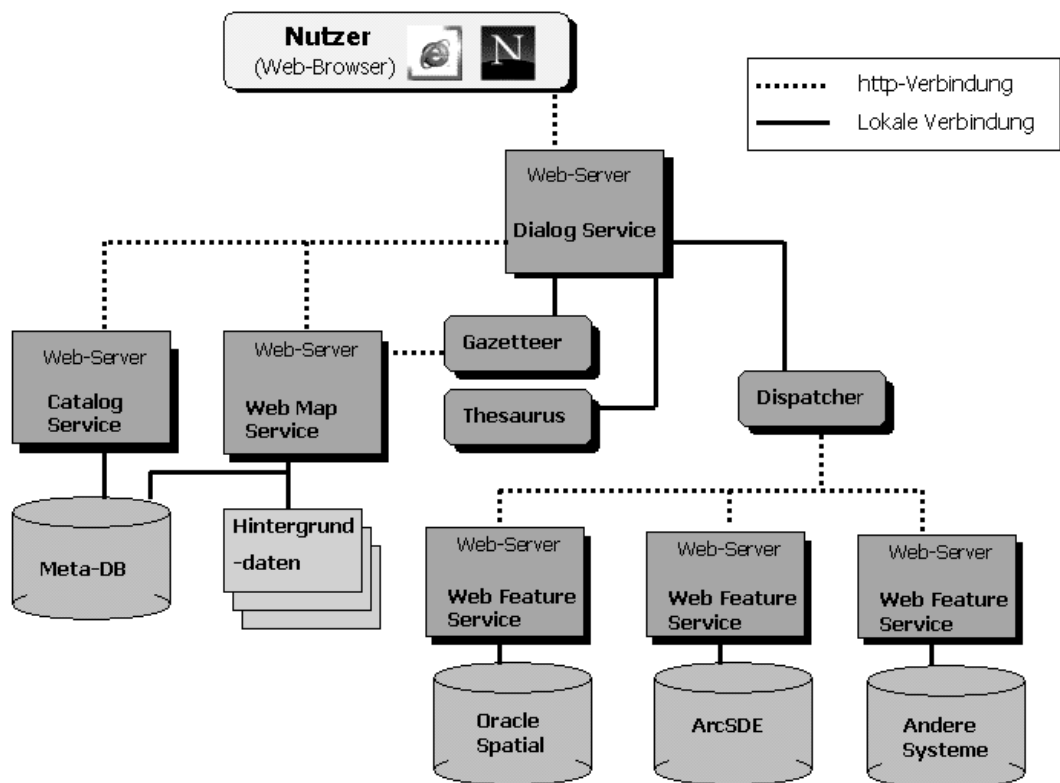


Abbildung 22. MISLux Architekturdiagramm

6.3.4 Systemarchitektur

Die Architektur von MISLux berücksichtigt mehrere Ebenen, die in Abbildung 22. dargestellt sind. Gestrichelte Linien in diesem Diagramm stehen für http-Verbindungen, während durchgezogene Linien andere Netz-

³⁷ Klasse dataset

³⁸ Klasse dataseries

³⁹ Klasse responsibleParty

⁴⁰ Klasse attribute

⁴¹ Klasse keywords

oder lokale Verbindungen symbolisieren. Elemente, die durch gestrichelte Linien verbunden sind, können somit beliebig im Internet positioniert sein.

Auf unterster Ebene stehen die Fachinformationssysteme der einzelnen Verwaltungen, die ihre Datenhaltung entweder mittels ESRI-Produkten (ArcSDE), einer *Simple Features* konformen Geodatenbank (Oracle 8i/9i) oder in sonstigen Systemen, wie beispielsweise dem proprietären Datenverwaltungssystem der Katasterverwaltung Luxemburgs, halten. Die Schnittstelle eines Fachinformationssystems nach außen wird durch einen *Web Feature Service* (WFS) festgelegt (OGC 2001b).

Die mittlere Ebene in Abb. 22. umfasst das MISLux-Portal, über welches jegliche Kommunikation zwischen Nutzern und Fachinformationssystemen durchgeführt wird. Der *Dialog Service* dient hierbei als direkte Schnittstelle zum Nutzer, mittels *Java Server Pages (JSP)* werden html-Seiten an den Browser des Anfragenden gesendet. Zur Unterstützung der Recherche greift der *Dialog Service* auf einen OGC-konformen *Web Map Service* (WMS, OGC 2000), einen Thesaurus und einen *Gazetteer* zurück. Anfragen des Nutzers werden an den *Catalog Service* weitergegeben, der die Metadaten des gesamten Datenbestandes abfragt und Informationen über die Suchkriterien erfüllenden Datensätze zurückliefert. Der WMS erfüllt die Funktion, die raumbezogene Recherche in Zusammenarbeit mit dem *Gazetteer* zu unterstützen und Informationen zur räumlichen Lage der Datensätze zu liefern. Dies umfasst Vorschaubilder, die entweder einfache umgebende Rechtecke der Datensätze zusammen mit den Grenzen Luxemburgs zeigen oder mittels eines Generalisierungsalgorithmus erzeugte *Thumbnail*-Bilder. Als Kataloginformation wird also eine grafische, prägnante Darstellungsform zusätzlich zur textuellen benutzt. Für diese Aufgaben greift der WMS auf Hintergrunddaten wie die Grenzen der Verwaltungseinheiten Luxemburgs zurück. Wird ein Datensatz identifiziert und angefordert, verteilt der *Dispatcher* die Anfrage an den betreffenden WFS.

Auf oberster Ebene steht der Client des MISLux-Nutzers, welcher mit Hilfe eines gängigen Web-Browsers dargestellt wird.

Die MISLUX-Architektur entspricht somit vollständig dem Referenzmodell für Katalogdienste, wie vom OGC spezifiziert (OGC 1999e) und in Abb. 7. dargestellt.

Die *Application* ist im Falle von MISLUX ein Catalog Service (CAS) Client, der im Prototyp mittels HTML, JavaScript und Java Server Pages (JSP) implementiert wurde. Ein solcher *thin* Client besitzt einen serverseitigen Teil und kann somit nicht komplett auf dem Rechner des Anfragenden ausgeführt werden. Im weiteren wird trotzdem die Gesamtheit des Clienten inklusive diesem serverseitigen Teil als MISLUX-CAS Client bezeichnet. Der in Abb. 22. erwähnte *Dialog Service* entspricht diesem serverseitigen Teil des MISLUX-CAS Client. Er nimmt die Eingaben des MISLUX-Nutzers entgegen und formuliert Anfragen mittels der *Filter Encoding Specification* des OGC (OGC 2001b). Diese Anfragen werden über die *Catalog Interfaces* an den *Application Server* weitergegeben. Als Kommunikationsprotokoll wird http verwendet, die Schnittstelle basiert, wie schon erwähnt, auf dem *Web Services Stateless Catalog Profile* (OGC 2001c). Der MISLUX-CAS Client unterstützt die Angabe räumlicher Suchparameter mittels einer interaktiven Karte, er ist hierfür mit einem WMS-kompatiblen (OGC 2000) *Web Map Server* Clienten gekoppelt. Im Prototyp wurde dieser mittels der *Open Source* Bibliothek JaGo umgesetzt⁴², es kann jedoch ein beliebiger WMS zum Einsatz kommen, solange er die *Web Map Server Specification* des OGC unterstützt. Die Angabe räumlicher Suchparameter wird außerdem mittels eines rudimentären Raumbezugsdienstes unterstützt, der indirekte auf direkte Georeferenzen abzubilden imstande ist.

Der *Application Server* entspricht dem serverseitigen Teil des MISLUX-Portales in Abb. 22. Dieser nimmt mittels der OGC *Catalog Interfaces* Anfragen entgegen und beantwortet diese. Zur Speicherung der Metadaten im *Metadata Store* wird im Prototypen eine Oracle Datenbank eingesetzt, die über eine *Spatial* Erweiterung verfügt. Mittels einer von dieser Datenbank zur Verfügung gestellten Simple Features Schnittstelle werden die Metadaten in

⁴² siehe <http://katla.giub.uni-bonn.de/jago>

einem ISO 19115 konformen Datenbankschema gespeichert⁴³. Zum Im- und Export der Daten wird XML eingesetzt.

Der Zugriff auf die verteilt liegenden Geodaten (*Data Store*) erfolgt mittels der *Web Feature Server* (WFS) Schnittstelle (OGC 2001b). Der zentrale *Application Server* verteilt die bei ihm eintreffenden *Access*-Anfragen an die entsprechenden Fachinformationssysteme, die hinter WFS-Schnittstellen verborgen sind.

Die Architektur von MISLux zeigt die konsequente Umsetzung der OGC Konzepte und Spezifikationen in einem konkreten Projekt. Durch konsequente Umsetzung internationaler Standards und Einsatz von Komponententechnologie entsteht ein offenes und erweiterbares Informationssystem, welches die Nachhaltigkeit der getätigten Investitionen gewährleistet.

6.4 Die Zukunft der raumbezogenen Metainformationsverarbeitung

Die zunehmende Standardisierung im technischen und inhaltlichen Bereich wird dazu führen, dass interoperable Geodatenkataloge und Geodienstekataloge in den nächsten Jahren Wirklichkeit werden. Diese Kataloge werden Teile von umfassenderen Systemen oder Strukturen sein, den sogenannten Geodaten-Infrastrukturen. Dieser Begriff bezeichnet das Konzept einer verlässlichen, unterstützenden Umgebung, die den Zugang zu raumbezogenen Informationen erlaubt und auf einer möglichst kleinen Menge von Spezifikationen, Protokollen und Vorgehensweisen aufbaut (NEBERT 2001). Eine *Spatial Data Infrastructure* (SDI) besteht aus der notwendigen Basistechnologie, aber auch der organisatorischen Umgebung, die ihre Existenz erst möglich macht. Aus technischer Sicht umfasst eine SDI Geodaten, Metainformationen, Möglichkeiten zu Datenentdeckung, ihrer Visualisierung und Evaluierung und einer Möglichkeit Zugriff, auf die Daten zu erhalten (NEBERT 2001). Raumbezogene Metainformationsverarbeitung nimmt eine zentrale Rolle in der Entwicklung von SDIs ein.

⁴³ Entspricht den *Internal Interfaces* e.g. *OGC Features*

Aufbauend auf den Standards von OGC, ISO und sich als de-facto Standards durchsetzenden Technologien wie XML werden sich Geodaten-Infrastrukturen auf verschiedenen administrativen Ebenen bilden. International gesehen existieren Initiativen dieser Art in einer Vielzahl von Ländern⁴⁴. In Deutschland stellt die GDI NRW einen Ansatz in dieser Richtung dar. Innerhalb dieser Initiative werden mit Hilfe der OGC-Spezifikationen Dienste zur Katalogisierung, Darstellung und zum Zugriff auf Geoinformationen entwickelt und im operativen Einsatz auf Interoperabilität getestet. Die Erkenntnisse und Lösungen dieser Initiative werden nicht nur im Ursprungsland Nordrhein-Westfalen, sondern auch darüber hinaus von Bedeutung sein.

GeoMIS.Bund stellt in erster Hinsicht einen Geodatenkatalog für Geodaten der öffentlichen Verwaltung Deutschlands dar. Daneben ist GeoMIS.Bund aber auch der erste Schritt zu einer Geodateninfrastruktur Deutschland (IMAGI 2001). GDI NRW und GeoMIS.Bund stellen voraussichtlich zwei Bausteine dar, aus denen sich eine deutsche Geodateninfrastruktur entwickeln wird. In welchem Zeitraum das geschieht und welche weiteren Bausteine daran teilhaben werden, wird sich in der näheren Zukunft zeigen.

Lokale und staatliche Infrastrukturen werden sich mehr und mehr zusammenschließen, was die Vision einer globalen Geodateninfrastruktur in erreichbare Nähe rücken lässt. Durch Ansätze, die geografische Informationstechnologie in Standard-IT Verfahren integrierbar machen, wird der Raumbezug von Daten über die klassischen GIS-Anwendungsgebiete hinaus an Bedeutung gewinnen. Diese integrierenden Ansätze sind vor allem in Bereichen der indirekten Georeferenzierung und ihrer automatisierten Verarbeitung zu finden. Metainformationsverarbeitung als integrierende Komponente wird eine Schlüsselrolle bei dieser Entwicklung spielen. Letztendlich wird der Aufbau einer dezentralen globalen Geodaten-Infrastruktur nur ein Schritt in der Entwicklung einer allgemeinen und globalen Informationsinfrastruktur sein.

⁴⁴ siehe beispielsweise MASSER (1999)

7 Zusammenfassung

Die raumbezogene Metainformationsverarbeitung spielt eine fundamentale Rolle bei der Konzeption und Entwicklung von verteilten Geografischen Informationssystemen, Umweltinformationssystemen und Geodaten-Infrastrukturen. Metainformationen erlauben die Identifizierung von Informationsressourcen und werden bei der Verarbeitung von Geoinformationen benötigt, indem sie den Entstehungskontext der die Daten kodierenden Daten beschreiben.

Die Terminologie und Definition der mit raumbezogener Informationsverarbeitung befassten Metainformation ist bislang pragmatisch motiviert und oftmals nicht zu verallgemeinern, wodurch bei der Planung komplexer GI-Projekte immer wieder Kommunikationsproblemen entstehen. Insbesondere die Unterscheidung zwischen Metadaten und „normalen“ Daten wird oftmals zum Anstoß langwieriger Diskussionen.

Diese Arbeit entwickelt einen Entwurf zur Lösung dieses Problems. Einen ingenieurswissenschaftlichen Referenzrahmen der Studie stellt das konzeptionelle Modell des *OpenGIS Consortiums* (OGC) dar. Das in dieser Arbeit entwickelte Modell kann als Erweiterung des abstrakten *OpenGIS Essential Models* angesehen werden. Es werden gängige Ansätze aus Forschung und Praxis zur Metainformationsverarbeitung vorgestellt, um dann vor diesem Hintergrund ein grundlegendes Modell der raumbezogenen Metainformationsverarbeitung abzuleiten. Als theoretischer Grundbau werden in der Semiotik entwickelte Konzepte benutzt und somit eine Teiltheorie der Semiotik entworfen. Zentral hierfür stehen die Begriffe Zeichen und Kode. Zeichen sind Informationsträger, sie erlauben die Weitergabe von Informationen an einen Empfänger, solange dieser den zur Entschlüsselung benötigten Kode besitzt. Die Interpretation und Anwendung der semiotischen Begrifflichkeit erlaubt eine funktionale Definition von Metainformationen und eine Unterteilung von Metainformationen in die Rollen Struktur-, Kontext- und Kataloginformationen. Diese Rollen entsprechen Konzepten der Semiotik und können anhand dieser Konzepte eindeutig identifiziert werden.

Strukturinformationen sind aus der technischen Speicherung von Daten abgeleitete Informationen, die Gegenstand klassischer *Data Dictionaries* sind, wie sie für Datenbanken benutzt werden. Sie entsprechen dem Kode, den eine Maschine benötigt, um Daten zu interpretieren. Kontextinformationen umfassen den Kode, den ein Mensch benötigt um das grafische Zeichen digitale Karte zu interpretieren. Kataloginformationen bilden den Grundbaustein von Zeichen, die Geodatenätze referenzieren, sind demnach Zeichen für Zeichen der Ausprägung Digitale Karte. Sowohl Originaldaten, Strukturdaten, als auch Kontextdaten können durch Anwendung in einem Katalog zu Kataloginformationen werden. Metainformationen, die zur Dokumentation von Diensten verwendet werden, werden als eine Erweiterung von Datensatz-bezogenen Metainformationen interpretiert, da Zugriff auf Datensätze als ein rudimentärer Dienst angesehen werden kann. Diese funktionalen Definitionen dienen als Grundlage der Unterscheidung zwischen Daten und Metadaten, beziehungsweise Informationen und Metainformationen. Eine Information wird zur Metainformation, wenn sie in einer der drei genannten Funktionen als Strukturinformation, Kontextinformation oder Kataloginformation benutzt wird. Die Unterscheidung der verschiedenen Arten von Information erfolgt aufbauend auf der Funktion, die sie erfüllt und ist somit kontextabhängig, was über den Rollenbegriff verdeutlicht werden soll.

Kataloginformationen werden innerhalb von Geodatenkatalogen durch Katalogeinträge dargestellt und besitzen Raumbezug. Da kartografische Darstellungsformen gegenüber textuellen den Vorteil einer stärkeren Prägnanz besitzen, wird die Empfehlung ausgesprochen, raumbezogene Metainformation mittels Karten darzustellen. Dies stellt ein weiteres Ergebnis dieser Arbeit dar.

Abschließend wird die Bedeutung eines solchen allgemeinen Modells der raumbezogenen Metainformationsverarbeitung diskutiert und gegenwärtige Entwicklungen vorgestellt. Anhand eines konkreten Beispiels, des Metainformationssystems für Geodaten des Großherzogtums Luxemburg (MISLux), werden die entwickelten Konzepte veranschaulicht. MISLux stellt ein modernes Katalogsystem für Geodaten dar und basiert auf den Konzepten

und Spezifikationen des OGC. Die konzeptionellen und theoretischen Überlegungen der Arbeit finden anhand dieses Projektes eine Umsetzung und Erläuterung. Ein Ausblick auf die zukünftigen Tendenzen der Geo-Metainformationsverarbeitung rundet die Arbeit ab. Ein Hauptaugenmerk wird hier auf die Entwicklung von Geodaten-Infrastrukturen und die Rolle, die hierbei der raumbezogenen Metainformationsverarbeitung zukommt, gelegt.

8 Anhang: Notation verwendeter UML-Diagramme

In dieser Arbeit werden ausschließlich Klassendiagramme in UML-Notation⁴⁵ verwendet. Die hierbei auftauchenden Konstrukte sind Klassen, Assoziationen, Rollen, Generalisierung/Spezialisierung, Aggregation und Pakete. Alle Definitionen sind angelehnt an OESTEREICH (1998).

8.1 Klasse



Abbildung 23. Klasse

Anstelle des Begriffes Klasse wird auch der Ausdruck Typ verwendet. Eine Klasse enthält eine Beschreibung der Struktur und des Verhaltens von Objekten, sie definiert folglich Attribute, Operationen und Semantik von Objekten. Alle Objekte einer Klasse entsprechen dieser Definition. Klassen werden durch Rechtecke dargestellt, die den Namen der Klasse tragen und zusätzlich die Namen von Attributen und Operationen beinhalten können. Die drei Rubriken Klassenname, Attribute, Operationen werden durch eine horizontale Linie getrennt.

8.2 Assoziation

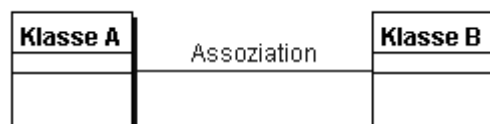


Abbildung 24. Assoziation

⁴⁵ alle Diagramme aufbauend auf UML Version 1.3

Eine Assoziation ist eine Verbindung zwischen Klassen und beschreibt somit die gemeinsame Semantik und Struktur einer Menge von Objektverbindungen. Assoziationen werden als eine Linie zwischen den beteiligten Klassen dargestellt. Assoziationen können mit Namen belegt werden, die dann entlang der Assoziation verzeichnet werden.

8.3 Rolle

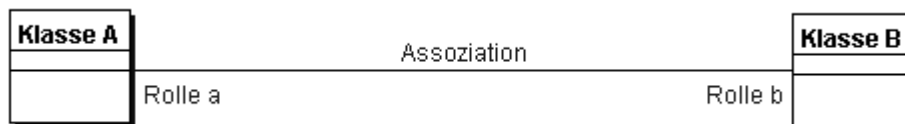


Abbildung 25. Rollen

Falls eine Beziehung in eine bestimmte Richtung gelesen werden kann, ist es möglich, mittels eines Rollennamens die Rolle des Zielobjektes in Beziehung zum Quellobjekt darzustellen. Der Rollenname beschreibt, wie das Objekt durch das in der Assoziation gegenüberliegende Objekt gesehen wird. Rollennamen werden entlang von Assoziationen dargestellt, wobei der jeweilige Rollenname näher zur zugehörigen Klasse dargestellt wird.

8.4 Generalisierung/Spezialisierung

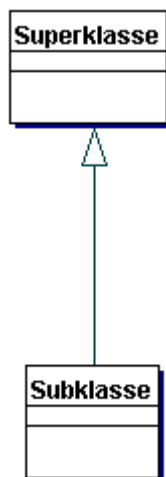


Abbildung 26. Generalisierung/Spezialisierung

Generalisierung und Spezialisierung sind Abstraktionskonstrukte, welche die hierarchische Gliederung der Semantik eines Modells erlauben. Hierbei werden Eigenschaften von der übergeordneten Klasse an die untergeordnete weitergegeben. Spezialisierung und Generalisierung stellen hierbei inverse Beziehungen dar, Spezialisierung ist die Aufgliederung einer Superklasse in Subklassen, Generalisierung die Zusammenfassung von Subklassen zu einer Superklasse. Vererbungsbeziehungen werden mit einem großen, nicht ausgefüllten Pfeil dargestellt, der von der Unterklasse zur Oberklasse zeigt.

8.5 Aggregation

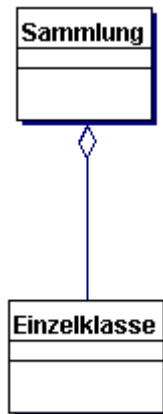


Abbildung 27. Aggregation

Aggregationen sind eine spezielle Art von Assoziation. Sie definieren eine Teil-von-Beziehung zwischen einer Sammlung von Klassen und einer Einzelklasse. Hierbei dient die Instanz der Sammlungsklasse als Behälter für die Objekte der Einzelklasse. Hierbei übernimmt die Sammlungsklasse stellvertretend Aufgaben für seine Teile. Im Gegensatz zu einer Assoziation sind die beiden beteiligten Klassen folglich nicht gleichberechtigt, sondern eine Klasse übernimmt die Verantwortung und Führung. Aggregationen werden auf gleiche Weise wie Assoziationen dargestellt, wobei die Sammlungsklasse durch eine Raute identifiziert wird.

8.6 Pakete

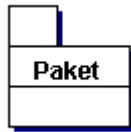


Abbildung 28. Paket

Pakete dienen der Gliederung von großen, umfangreichen Modellen. Sie sind Sammlungen von anderen Modellelementen beliebigen Typs, wie beispielsweise Klassen oder andere Pakete. Pakete werden aufgrund logischer Zusammenhänge gebildet.

9 Literatur

- Abler, R., Adams, J. & Gould, P. (1971): *Spatial Organization: The Geographer's View of the World*. Englewood Cliffs, N.J., USA
- [ANSI] American National Standards Institute (Hrsg.) (1997a): *American National Standard for Information Systems – Spatial Data Transfer Standard. Part 1: Logical Specifications*. Internet:
ftp://sdts.er.usgs.gov/pub/sdts/standard/latest_draft/pdf/part1.pdf
- [ANSI] American National Standards Institute (Hrsg.) (1997a): *American National Standard for Information Systems – Spatial Data Transfer Standard. Part 2: Spatial Features*. Internet:
ftp://sdts.er.usgs.gov/pub/sdts/standard/latest_draft/pdf/part2.pdf
- [ANSI] American National Standards Institute (Hrsg.) (1997a): *American National Standard for Information Systems – Spatial Data Transfer Standard. Part 3: ISO 8211 Encoding*. Internet:
ftp://sdts.er.usgs.gov/pub/sdts/standard/latest_draft/pdf/part3.pdf
- Arctur, D., Hair, D., Timson, G., Martin, E. & Fegeas, R. (1998): *Issues and prospects for the next generation of the spatial data transfer standard (SDTS)*. In: *Int. J. Geographical Information Science* 12(4), S. 403-425.
- Arndt, H.-K., Günther, O. & Pluskat, H. (2001): *Wissensmanagement für das Umweltmanagement*. In: *Tochtermann, K. & Riekert, W.-F. (2001): Neue Methoden für das Wissensmanagement im Umweltschutz, Marburg, S. 211-219.*
- Aronoff, S. (1991): *Geographic Information Systems: A Management Perspective*. Ottawa.
- Augstein, B. & Greve, K. (1994): *Umweltanwendungen Geographischer Informationssysteme - Bausteine für einen Referenzrahmen*. In: *Dollinger, F. und Strobl, J. (Hrsg.) (1994): Angewandte Geographische Informationstechnologie VI. Beiträge zum GIS-Symposium 6.-8. Juli 1994, Salzburg, S. 19-28.*
- Balfanz, D. & Göbel, S. (1999): *Bridging geospatial metadata standards towards distributed metadata information systems*. In: *IEEE Metadata*

- proceedings 1999. Internet:
<http://computer.org/conferen/proceed/meta/1999/papers/40/dbalfanz.html>.
- Bandholtz, T., Bös, R. & Rüter, M. (2000): *The German Environmental Information Network*. In: Cremers, A. & Greve, K. (2000): Computer Science for Environmental Protection '00. Umweltinformation für Planung, Politik und Öffentlichkeit. Marburg. S. 632-645.
- Bell, S. (1997): *Entdeckung von Metadaten zur semantischen Anfrageoptimierung in relationalen Datenbanken*. Aachen.
- Bertin, J. (1974): *Graphische Semiologie. Diagramme, Netze, Karten*. Berlin.
- Bilo, M. & Streuff, H. (2000): *Das Umweltinformationsnetz Deutschland GEIN 2000 - Fachliche Anforderungen an ein Forschungs- und Entwicklungsvorhaben*. In: Tochtermann, K. u. Riekert, W.-F. (Hrsg.) (2000): Hypermedia im Umweltschutz, 3. Workshop, Ulm, Marburg (Umwelt-Informatik aktuell, Bd. 24), S. 217–228.
- Bishr, Y. (1998): *Overcoming the semantic and other barriers to GIS interoperability*. In: Int. J. Geographical Information Science, 12(4), S. 299-314.
- Bray, T. (2001): *What is RDF?*. Internet:
<http://www.xml.com/pub/a/2001/01/24/rdf.html>
- Bretherton, F. (1994): *Reference Model for Metadata: A Strawman*. In: IEEE Computer Society Technology Commission on MSS Metadata Workshop, February 1994, University of Texas, Austin. Internet:
http://www.llnl.gov/liv_comp/metadata/papers/whitepaper-bretherton.html.
- Broy, M. (1992): *Informatik. Eine grundlegende Einführung. Teil 1*. Berlin.
- Buehler, K. & McKee, L. (Hrsg.) (1998): *The OpenGIS Guide. - Introduction to Interoperable Geoprocessing and the OpenGIS Specification*. Draft - June 3, 1998 - Third Edition. Wayland, MA, USA.
- [CEN] Comité Européen de Normalisation (Hrsg.) (1998): ENV 12657 *Geographic Information – Data description – Metadata*. Internet:
http://comelec.afnor.fr/servlet/ServletForum?form_name=cForumPage

&file_name=Z13C%2FPUBLIC%2FWEB%2FENGLISH%2Fpren.htm&login=invite&password=invite.

Cleveland, G. (1998): *Digital Libraries: Definitions, Issues and Challenges*.

UDT Occasional Paper #8. Internet:

<http://www.ifla.org/VI/5/op/udtop8/udtop8.htm>.

Consorti, F., Merialdo, P. & Sindoni, G. (1996): *Metadata Reference Model for Medical Documentation: a Hypermedia Proposal*. In: Proceedings of the First IEEE Metadata Conference. Internet:

<http://www.computer.org/conferences/meta96/sindoni/ieee.html>.

Couclelis, H. (1992): *People manipulate Objects (but cultivate Fields):*

Beyond the Raster-Vector Debate in GIS. In: Frank, A.U., Campari, I.

und Formentini, U. (Hrsg.) (1992): *Theories and Methods of Spatio-Temporal Reasoning in Geographic Space*, Proceedings, Berlin, S. 65-77.

Danko, D. M. (1997): *Perspectives in the development of ISO metadata standards*. Internet:

<http://fgdc.er.usgs.gov/publications/documents/metadata/nimapaper.html>

[DCME] Dublin Core Metadata Initiative (Hrsg.) (2001): *Dublin Core Metadata Initiative Overview*. Internet:

<http://www.dublincore.org/about/>

Denzer, R. & Güttler, R. (1994): *Integration von Umweltdaten*. In: Page, B., Hilty, H. (Hrsg.) (1994): *Umweltinformatik. Informatikmethoden für Umweltschutz und Umweltforschung*. München-Wien, S. 227-240.

Eco, U. (1977): *Zeichen. Einführung in einen Begriff und seine Geschichte*. Frankfurt.

Eco, U. (1985): *Semiotik und Philosophie der Sprache*. In: Henschen, H.-H. (Hrsg.) (1985): *Supplemente, Band 4*. München.

Eco, U. (1987): *Semiotik. Entwurf einer Theorie der Zeichen*. In: Henschen, H.-H. (Hrsg.) (1987): *Supplemente, Band 5*. München.

[ESRI] Environmental Systems Research Institute (Hrsg.) (2001): *ESRI Profile of the Content Standard for Digital Geospatial Metadata*. Internet:

http://arconline.esri.com/arconline/whitepapers/ao_/J8709_Geospatial_Metadata_Profile.pdf?PID=1.

- Fitzke, J. (2001): *Verwaltung von Raumbezügen in Informationssystemen für den Umwelt- und Naturschutz*. In: Proceedings of the 15th International Symposium Informatics for Environmental Protection. Marburg, S. 897-902
- Fitzke, J., Greve, K. Lessing, R. & Müller, M. (2000): *GIS-Pool: eine Managementkomponente für Geodaten in komplexen Organisationen*. In: Strobl, J., Blaschke, T. & Griesebner, G. (Hrsg.) (2000): *Angewandte Geographische Informationsverarbeitung XII*, Salzburg, S. 144-150.
- [FGDC] Federal Geographic Data Committee (Hrsg.) (1994): *Content Standards for Digital Geospatial Metadata*. U.S. Government, Federal Geographic Data Committee, Washington, DC. Internet: <http://www.fgdc.gov/>
- Fonseca, F. (2000): *GIS_ontology.com*. In: GIScience 2000, First International Conference on Geographical Information Science. Internet : <http://www.giscience.org/GIScience2000/papers/218-Fonseca.pdf>
- Fornefeld, M. & Oefinger, P. (2001): *Marktstudie: Aktivierung des Geodatenmarktes in Nordrhein-Westfalen*. Internet: <http://www.media.nrw.de/imperia/md/content/publikatio/47.pdf>.
- Frank, A. (1996): *Der Nutzen und der Preis von geographischer Information*. In: Dollinger, F. & Strobl, J. (Hrsg.) (1996): *Angewandte geographische Informationsverarbeitung: Beiträge zum AGIT-Symposium Salzburg `96*. Salzburg, S. 61-70. Internet: <http://www.sbg.ac.at/geo/agit/papers96/frank.htm>
- Fretter, K. (2000): *Teil-automatisierte Erfassung von geographischen Metainformationen aus digitalen Datenbeständen*. In: Strobl, J., Blaschke, T., und Griesebner, G. (Hrsg.) (2000): *Angewandte Geographische Informationsverarbeitung XII*, Salzburg, S. 159-166.
- Goodchild, M., Egenhofer, M., Kemp, K., Mark, D. & Sheppard, R. (1999): *Introduction to the Varenius project*. In: *Int. J. Geographical Information Science* 13(8), S. 731-745.

- Greve, K. (1996): *Objekte der Metadatenverarbeitung im Umweltbereich*. In: Datenqualität und Metainformation in Geo-Informationssystemen. Tagungsband zum Workshop vom 7. bis 8. Oktober 1996 an der Universität Rostock. Rostock.
- Greve, K. (1999): *Ein Referenzmodell zur Beschreibung der Konzeption von UIS*; In: Bock, M., Greve, K. & Kuhn, W. (Hrsg.) (1999): Offene Umweltinformationssysteme- Chancen und Möglichkeiten der OpenGIS- Entwicklung im Umweltbereich, Münster, S. 33- 45.
- Greve, K. & Häuslein, A. (1994): Metainformationen in Umweltinformationssystemen. In: Hilty, L.M., Jaeschke, B., Page, B. & Schwabl, A.(Hrsg) (1994): Informatik für den Umweltschutz. 8. Symposium, Hamburg, S.169-178.
- Greve, K., Lessing, R. & Müller, M. (2001): *Operationell nutzbare Metainformationen zur Beschreibung von Raum- und Zeitbezug*. In: Strobl, J., Blaschke, T. & Griesebner, G. (Hrsg.) (2001): Angewandte Geographische Informationsverarbeitung XIII, Salzburg, S. 209-214.
- Günther, O. (1998): *Environmental Information Systems*. Berlin.
- Günther, O., Lessing, H. & Swoboda, W. (1996): *UDK: A European Environmental Data Catalogue*. In: Proceedings of the Third International Conference/Workshop on Integrating GIS and Environmental Modelling, Santa Fe, New Mexico, USA. Internet: http://ncgia.ncgia.ucsb.edu:80/conf/SANTA_FE_CD-ROM/sf_papers/guenther_oliver/my_paper.html.
- Günther, O. & Voisard, A. (1997): *Metadata in Geographic and Environmental Data Management*. In Klas, W. & Sheth, A. (Hrsg.) (1997): Managing Multimedia Data: Using Metadata to Integrate and Apply Digital Data. Columbus, S. 57-87.
- Hake, G., Grünreich, D. & Meng, L. (2002): *Kartographie: Visualisierung raum-zeitlicher Informationen*. Berlin.
- Hakimpour, F. & Timpf, S. (2001): *Using Ontologies for Resolution of Semantic Heterogenity in GIS*. Paper presented at the 4th AGILE Conference on Geographic Information Science 2001. Internet: <http://www.ifi.unizh.ch/dbtg/Projects/MIGI/publication/agile2001.pdf>

- Hallmark, G. (1995): *The Oracle Warehouse*. In: Proceedings of the 21st. Conference on Very Large Data Bases 1995. Internet: <http://www.vldb.org/conf/1995/P707.PDF>.
- Heuer, A. & Saake, G. (1997): *Datenbanken: Konzepte und Sprachen*. Bonn
- Holman, G. (2000): *What is XSLT*. Internet: <http://www.xml.com/pub/a/2000/08/holman/index.html>
- [IMAGI] Interministerieller Ausschuss für Geoinformation (Hrsg.) (2000): *Konzeption eines effizienten Geodatenmanagement des Bundes*. Frankfurt am Main. Internet: <http://www.imagi.de/>.
- [IMAGI] Interministerieller Ausschuss für Geoinformation (Hrsg.) (2001): *Beschlüsse zur 7. Sitzung des Interministeriellen Ausschuss für Geoinformationswesen (IMAGI) am 10. Oktober 2001. TOP 4: Strategie für den Aufbau einer Geodateninfrastruktur Deutschland (GDI.de)*. Internet: <http://www.imagi.de/Beschlüsse.htm>.
- InGeoForum (Hrsg.) (2000): *InGeo-Metadatenformat (InGeo-MDF)*. Version 2.1. Internet: <http://www.ingeoic.de/>.
- [ISO] Norwegian Technology Standards Institution (NTS) (Hrsg.) (2000): DIS 19115, *Geographic Information – Metadata*. Internet: <http://www.isotc211.org/>.
- [ISO] Norwegian Technology Standards Institution (NTS) (Hrsg.) (2001): 19119, *Geographic Information – Services*. Internet: <http://www.isotc211.org/>.
- [IuK] IuK Initiative Information und Kommunikation der wissenschaftlichen Fachgesellschaften Deutschlands (Hrsg.) (2000): *Information, Knowledge and Knowledge Management*. Proceedings of the Annual Meeting of the “IuK-Initiative der wissenschaftlichen Fachgesellschaften”. Internet: <http://www.soziologie.de/iuk2000/>.
- Krasemann, H. (1996): *Meta-Information – Was ist das? Ein Resumee des Workshops „Umweltdaten verstehen durch Metainformation“ des Arbeitskreises Umweltdatenbanken der Fachgruppe 4.6. – Informatik im Umweltschutz der GI von 17. Bis 19. Mai in Lauenburg und am GKSS-Forschungszentrum Geesthacht*. In: Kremers, H. & Krasemann,

- H. (Hrsg.) (1996): *Umweltdaten verstehen durch Metainformation (Praxis der Umweltinformatik 6)*. Marburg, S.
- Kriz, J., Lück, H. & Heidbrink, H. (1990): *Wissenschafts- und Erkenntnistheorie. Eine Einführung für Psychologen und Humanwissenschaftler*, 2. Aufl. Opladen.
- Kruse, F. & Eichler, M. (2001): *Der Windows UDK 4.3 auf Basis der SOAP-Technologie*. In: Proceedings of the 15th International Symposium Informatics for Environmental Protection 2001. Marburg, S. 915-920.
- Kuhn, W. (1997). *Toward Implemented Geoprocessing Standards: Converging Standardization Tracks for ISO/TC 211 and OGC*, White Paper. ISO/TC 211-N418. Internet:
http://www.opengis.org/techno/white_papers/211n418.pdf.
- Lessing, H. & Weiland, H. (1990): *Der Umwelt-Daten-Katalog Niedersachsen. Konzept, Stand der Anwendung und Fortentwicklung*. In: Pillmann, W., et al. (Hrsg.) (1990): *Informatik für den Umweltschutz*. Marburg, S. 402-410.
- Lethen, J. (1999a): *ETC/CDS – das Europäische Themenzentrum Datenquellenkatalog im Niedersächsischen Umweltministerium: Zugang zu Umweltinformationen. Genereller Überblick über das Tätigkeitsfeld des ETC/CDS*. Internet:
http://www.mu.niedersachsen.de/cds/etc-cds_neu/library/etc-faltblatt.pdf.
- Lethen, J. & Jensen, S. (1999b): *General Information on the European Topic Centre on Catalogue of Data Sources. People, purposes, tasks and products*. Internet: http://www.mu.niedersachsen.de/cds/etc-cds_neu/library/ETC_general_info.pdf
- Malhotra, Y. (1998): *Knowledge Management, Knowledge Organizations & Knowledge Workers: A View from the Front Lines*. Internet:
<http://www.brint.com/interview/maeil.htm>.
- Malhotra, Y. (2002): *A Case For Knowledge Management: Rethinking Management for the New World of Uncertainty and Risk*. Internet:
<http://www.kmbook.com/>

- Mark, D., Freksa, C., Hirtle, S., Lloyd, R. & Tversky, B. (1998): *Cognitive models of geographical space*. In: Int. J. Geographical Information Science. 13(8), S. 747-774.
- Masser, I (1999): *All shapes and sizes: the first generation of national spatial data infrastructures*. In: Int. J. Geographical Information Science 13(1), S. 67-84.
- Minnesota Historical Society (Hrsg.) (2000): *Data warehouse description*.
Internet: <http://www.mnhs.org/preserve/records/dwintro.html>
- Müller, M. (2001a): *Geografische Metadaten: Zur Bedeutung nationaler und internationaler Standardisierungsinitiativen*. In: Neue Methoden für das Wissensmanagement im Umweltschutz. 4. Workshop des GI-Arbeitskreises Hypermedia im Umweltschutz und Workshop 3 der GI-Initiative Environmental Markup Language 2001, Marburg, S. 179-184.
- Müller, M. (2001b): *Entwurf einer Theorie umweltbezogener Metainformationsverarbeitung*. In: Proceedings of the 15th International Symposium Informatics for Environmental Protection 2001, Marburg, S. 927-932.
- Nebert, D. (Hrsg.) (2001): *The SDI cookbook*. Version 1.1. Internet:
<http://www.gsdi.org>.
- Niedersächsisches Umweltministerium (Hrsg.) (1998): *UDK 4.0 Datenmodell*.
Internet: <http://www.mu.niedersachsen.de/udk/>.
- Nikolai, R., Kazakos, W., Kramer, R., Behrens, S., Swoboda, W. & Kruse, F. (1999): *WWW-UDK 4.0: Die neue Generation eines Web-Portals zu deutschen und österreichischen Umweltdaten*. In: Tagungsband des 13. Internationalen Symposiums "Informatik für den Umweltschutz 1999, Magdeburg, S. 347-361.
- Oestereich, B. (1998): *Objektorientierte Softwareentwicklung: Analyse und Design mit der Unified modeling language*. - 4. Auflage, München; Wien.
- [OGC] OpenGIS Consortium (Hrsg.) (1999a): *Abstract Specification Topic 5 - Features (Version 4)*. OpenGIS Project Document 99-105r2. Internet:
<http://www.opengis.org/techno/abstract.htm>.

- [OGC] OpenGIS Consortium (Hrsg.) (1999b): *Abstract Specification Topic 9 - Quality (Version 4)*. OpenGIS Project Document 99-109r1. Internet: <http://www.opengis.org/techno/abstract.htm>.
- [OGC] OpenGIS Consortium (Hrsg.) (1999c): *Abstract Specification Topic 11 - Metadata (Version 4)*. OpenGIS Project Document 99-111r1. Internet: <http://www.opengis.org/techno/abstract.htm>.
- [OGC] OpenGIS Consortium (Hrsg.) (1999d): *Abstract Specification Topic 13 - Catalogs (Version 4)*. OpenGIS Project Document 99-113. Internet: <http://www.opengis.org/techno/abstract.htm>.
- [OGC] OpenGIS Consortium (Hrsg.) (1999e): *Catalog Interface Implementation Specification (Version 1.0)*. OpenGIS Project Document 01-033. Internet: <http://www.opengis.org/techno/implementation.htm>.
- [OGC] OpenGIS Consortium (Hrsg.) (2000): *OpenGIS® Web Map Server Interface Implementation Specification Revision 1.0.0*. OpenGIS Project Document 00-028. Internet: <http://www.opengis.org/techno/implementation.htm>.
- [OGC] OpenGIS Consortium (Hrsg.) (2001a): *Abstract Specification Topic 12 – Service Architecture (Version 4.2)*. OpenGIS Project Document 01-033. Internet: <http://www.opengis.org/techno/abstract.htm>.
- [OGC] OpenGIS Consortium (Hrsg.) (2001b): *OpenGIS® Filter Encoding Specification Version 0.0.6*. OpenGIS Project Document 01-033. Internet: <http://www.opengis.org/techno/rfc13info.htm>.
- [OGC] OpenGIS Consortium (Hrsg.) (2001b): *OpenGIS® Filter Encoding Specification Version 0.0.6*. OpenGIS Project Document 01-033. Internet: <http://www.opengis.org/techno/rfc13info.htm>.
- [OGC] OpenGIS Consortium (Hrsg.) (2001c): *Geography Markup Language(GML) 2.0, OpenGIS® Implementation Specification*. OpenGIS Project Document 01-029. Internet: <http://www.opengis.org/techno/implementation.htm>.
- Pick, T., Jensen S. & Menger, M. (2000): *Access to environmental information – EEAs WebCDS 2.6*. In: Tochtermann, K. & Riekert, W.-F. (Hrsg.)

- (2000): *Hypermedia im Umweltschutz*, 3. Workshop 2000, Ulm, Marburg (Umwelt-Informatik aktuell, Bd. 24), S. 51–54.
- Pick, T. & Saarenma, H. (2001): *The Dublin Core Element Set as a Reference Framework for Environmental Metadata*. In: Tochtermann, K. & Riekert, W.-F. (2001): *Neue Methoden für das Wissensmanagement im Umweltschutz*, Marburg, S. 211-219.
- Pick, T., Kvist, E. & Lounma, K.. (2001): *Environmental Metadata – Doing it the Dublin Core Way?*. In: *Proceedings of the 15th International Symposium Informatics for Environmental Protection 2001*. Marburg, S. 659-670.
- Reich, L. & Panagiotis, P. (Hrsg.) (2001): *Web Services Stateless Catalog Profile Version 0.06*. OpenGIS Project Document 01-062.
- Riekert, W.-F. & Treffler, P. (2000): *Georeferenzierung als Mittel zur Erschließung von Fachinformationen in Internet und Intranet*. In: Greve, K. & Cremers, A.B. (Hrsg)(2000): *Umweltinformatik '00 – Umweltinformation für Planung, Politik und Öffentlichkeit*, 12. int. Symp. "Informatik für den Umweltschutz" der Gesellschaft für Informatik, 2000, Bonn, Marburg (Umweltinformatik aktuell, Bd. 26), S. 413–421.
- Senatsverwaltung für Stadtentwicklung, Umweltschutz und Technologie (Hrsg.) (1996): *Umweltinformationssystem Berlin (UIS) der Senatsverwaltung für Stadtentwicklung, Umweltschutz und Technologie*. Internet:
<http://www.stadtentwicklung.berlin.de/umwelt/umweltatlas/>.
- Sompel, H. van de & Lagoze, C. (2001): *The Open Archives Initiative Protocol for Metadata Harvesting*. Internet:
<http://www.openarchives.org/OAI/openarchivesprotocol.htm>
- Stachowiak, H. (1983): *Systematischer Neopragmatismus: Grundriss einer zeitgenössischen Erkenntnistheorie*. In: Weingartner, P. & Czermak, H. (Hrsg) (1983): *Erkenntnis und Wissenschaftstheorie: Akten d. 7. Internat. Wittgenstein-Symposiums*, 22.-29. August 1982, Kirchberg am Wechsel (Österreich), Wien, S. 129-137.

- Strobl, J. (1995): *Grundzüge der Metadatenorganisation für GIS*. In:
Salzburger Geographische Materialien, Heft 22, S. 275-286.
- [UBA] Umweltbundesamt (Hrsg.) (2000a): *Pressemitteilung "Per Mausklick
in Deutschlands größtes Umwelt-Informationsnetzwerk"*.
Internet: [http://www.umweltbundesamt.de/uba-info-
presse/pressemitteilungen/p-2500-d.htm](http://www.umweltbundesamt.de/uba-info-
presse/pressemitteilungen/p-2500-d.htm)
- [UBA] Umweltbundesamt (Hrsg.) (2000b): *G2k-Profil*. Version 1.1. Internet:
<http://www.gein.de/2000/profile-11.htm>.
- Vlist, E. van der (2000): *Using W3C XML Schema*. Internet:
<http://www.xml.com/pub/a/2000/11/29/schemas/part1.html>
- Wahrig, G. (1999): *Fremdwörterlexikon*. Gütersloh.
- Worboys, M. (1995): *GIS - A Computing Perspective*, London.