

Netzzugang zu heterogenen, verteilten Umweltdaten

Björn Gehlsen¹, Ralf Kriebisch¹,
Hansjörg Krasemann², Winfried Lamersdorf¹, Bernd Page¹

Abstract

Connecting databases to the internet makes their contents widely available. Environmental data can also be reached, and there are many data collections accessible in this area. So it is easy to get confused about where to find and how to interpret relevant data. This paper describes the development of a user interface for heterogenous, distributed environmental data, enabling a comfortable and direct access to different data collections for potential users. A three-tier client server architecture is established, which is based on the open communication infrastructure CORBA and is ready to embed other data collections of related areas. Its search and navigation methods should help even incidental users to retrieve data with minimum effort.

1 Einleitung

Mit dem Internet steht heute zu moderaten Kosten eine Technologie zur Verfügung, die über angeschlossene Rechner eine weltweite Kommunikation erlaubt. WWW-Browser, die für praktisch jede gängige Rechnerplattform existieren, ermöglichen den Abruf von Dokumenten, die auf entfernten Rechnern bereitgestellt sind, und zeigen sie auf dem eigenen Bildschirm in aufbereiteter Form an.

Neben einzelnen Dokumenten sind auch Datenbanken über das Internet erreichbar. Nach der Erteilung entsprechender Zugriffsrechte können erfahrene Nutzer Abfragen an die Datenbank in der betreffenden Datenbankabfragesprache interaktiv oder innerhalb eigener Anwendungsprogramme vornehmen. Um auch weniger erfahrenen Nutzern einen Zugriff auf die Datenbank zu ermöglichen, wird meist eine komfortablere Schnittstelle in Form eines Browser-Formulars angeboten. In einzelne Formularfelder können Werte eingegeben werden, durch die sich vorformulierte Anfragen parametrisieren lassen.

Aus der Möglichkeit, eine bestimmte Datenbank komfortabel zu erreichen, erwächst natürlicherweise der Wunsch, auch auf andere Datenhaltungssysteme mit

¹ Universität Hamburg, Fachbereich Informatik, Vogt-Kölln-Straße 30, D-22527 Hamburg, {gehlsen, kriebisch, lamersdorf, page}@informatik.uni-hamburg.de

² GKSS-Forschungszentrum Geesthacht, Institut für Gewässerphysik, Max-Planck-Straße 1, D-21502 Geesthacht, krasemann@gkss.de

thematisch verwandtem Inhalt ebenfalls über eine einheitliche Benutzungsschnittstelle zugreifen zu können. Viele WWW-Anwendungen sind jedoch nicht flexibel genug ausgelegt, um auch Zugriffe zu anderen Datenhaltungssystemen zu ermöglichen. Einerseits sind die Rechneradressen von erreichbaren Datenbanken in CGI-Skripten fest vorgegeben, andererseits unterstützen die Schnittstellen dieser Skripte oft nur bestimmte Arten von Datenhaltungssystemen, z.B. nur relationale Datenbanksysteme eines Herstellers.

Das Auffinden gewünschter Dokumente oder Datenbestände stellt oft ein erhebliches Problem dar. Zum gegenwärtigen Zeitpunkt werden Dokumente in der Regel von Suchmaschinen erfaßt und indiziert. Dabei wird meist nur Freitext ohne wichtige Kontextinformation aufbereitet, da die notwendigen Metadaten nicht direkt im Dokument zu finden sind. Daten, die nur über Formulare zu erreichen sind, werden gar nicht erst von Suchmaschinen erfaßt. Entsprechend fallen die Ergebnisse von Suchanfragen aus: Wenn die Suchbegriffe nur ohne Kontextinformation benutzt werden können, so liefern Suchanfragen Ergebnislisten mit mehreren tausend Einträgen. Wird nach speziellen Markierungen gesucht, so sind nur wenige Dokumente vorhanden, die diese Markierungen aufweisen.

Das Suchproblem innerhalb von Datenhaltungssystemen wird noch dadurch verschärft, daß viele der vorgesehenen Dokumentationsfelder für Datensätze ungenutzt bleiben. Häufig ist der Nutzen dieser Dokumentationsarbeiten für diejenigen Personen, die Daten bereitstellen wollen, nicht direkt erkennbar. Erst wenn ein Recherche-system zur Verfügung steht, das umfangreiche Datenbestände erschließt, ausreichend vielen Benutzern zugänglich ist und komfortable Suchmöglichkeiten bietet, wird eine Rückmeldung und Anerkennung der geleisteten Arbeit erfolgen. Hieraus sollte wiederum die Motivation für den zusätzlich zu betreibenden Aufwand erwachsen.

Das Projekt TIDE¹ versucht hier Ansätze für mögliche Lösungen aufzuzeigen. Dazu wird eine Architektur für ein Zugriffssystem entworfen, die hinsichtlich heterogener Datenbestände erweiterungsfähig ist und auf modernen Kommunikationstechniken basiert. Weiterhin entsteht eine übersichtliche Oberfläche für komfortable Recherchemöglichkeiten. Datenhaltenden Stellen wird eine nachträgliche Dokumentation über spezielle Formularfelder ermöglicht, und sie werden auf die Wichtigkeit der Dokumentationsarbeit hingewiesen. Ohne diese Dokumentationsarbeit wäre eine Integration mit vertretbarem Aufwand nicht zu leisten.

¹ Tools for an Integrative View of Distributed Environmental Data: <http://tide-www.informatik.uni-hamburg.de>

2 Anforderungen an ein Zugriffssystem

Im Projekt TIDE besteht die vorrangige Aufgabe darin, eine Nutzerführungskomponente für eine bestimmte Datenbank (Wattenmeerdatenbank WADABA) zu erstellen. Dabei wurden auch die in der Einleitung erwähnten Probleme bei der Datensuche deutlich. Später entstand die Idee, auch andere Datenhaltungssysteme in die Gesamtarchitektur einzubinden. Nun soll ein Zugriffssystem erarbeitet werden, das mit Hilfe von WWW-Techniken den Zugriff auf wichtige Datenbestände aus dem Bereich Wattenmeer und küstennaher Gewässer von beliebigen Rechnern aus realisiert. Es soll gleichzeitig offen sein, um weitere Datenbestände verwandter Disziplinen zu erschließen. Als potentielle Nutzer eines derartigen Zugriffssystems ist besonders an Entscheidungsträger aus Politik, Wirtschaft und Verwaltung, Wissenschaftler verschiedenster Fachrichtungen und auch an interessierte Privatpersonen gedacht. In den folgenden Abschnitten werden die speziellen Anforderungen an ein solches System näher erläutert.

2.1 Plattformunabhängige Nutzungsschnittstelle

Die Erhebung von Umweltdaten ist mit hohen Kosten verbunden. Eine möglichst breitgefächerte Nutzung vorhandener Daten rechtfertigt somit den Aufwand für ihre Erhebung besonders. Um die verfügbaren Daten einem großen Anwenderkreis zugänglich zu machen, ist es erforderlich, den Zugang von vielen Rechnerplattformen aus zu realisieren. Als Mindestanforderung kann die Realisierung eines Zugriffs auf die Daten unter MS Windows, MacOS und wichtigen Unix-Varianten angesehen werden. Für diese Plattformen existieren Internet-Browser mit nahezu gleichen Benutzungseigenschaften, so daß bei Einsatz von WWW-Techniken die Heterogenität der Client-Hardware kein wesentliches Problem mehr darstellt.

2.2 Flexibilität zur Einbindung anderer Datenhaltungssysteme

Mit der Möglichkeit bequem vom Arbeitsplatz aus auf einen Datenbestand zugreifen zu können, ist der Wunsch absehbar, auch andere Datenbestände unter der gleichen Oberfläche zu erreichen. In einem weiteren Schritt sind Suchanfragen über die Grenzen von Datenbestände hinweg angebracht. Mit Blick auf spätere Erweiterungen soll das System die Flexibilität besitzen, Datenbestände zu integrieren, deren Existenz zur Entwicklungszeit noch nicht bekannt war, oder die erst in der Zukunft erhoben werden. Deshalb sind allgemeine Schnittstellen zu definieren, die es den Betreibern weiterer Datenbestände erlauben, ihre Datenbestände über das TIDE-System verfügbar zu machen. Die Implementation dieser Schnittstellen wird üblicherweise in

der Regie der datenhaltenden Institution durchgeführt, da nur dort das notwendige Wissen über die eingesetzten Datenstrukturen zur Verfügung steht.

2.3 Komfortable Such- und Navigationsmöglichkeiten

Die potentiellen Nutzer des Systems verfügen nur selten über detaillierte IT-Kenntnisse. Daher ist unter dem Aspekt einer breiten Nutzung einmal erhobener Daten die Benutzerfreundlichkeit ein wichtiges Kriterium für die Akzeptanz des Zugriffsystems. Es sollen Mechanismen sowohl für die gezielte Suche als auch für die Navigation zur Verfügung gestellt werden.

Unter Suche versteht man dabei den Abgleich der Metadaten mit den spezifizierten Kriterien einer Anfrage. Suchanfragen können die gewünschten Datenobjekte sehr genau eingrenzen, setzen aber ein Mindestmaß an Überblick über den verfügbaren Datenbestand voraus.

Demgegenüber bietet die Navigation einem unerfahrenen Nutzer die Möglichkeit, sich den Datenbestand Schritt für Schritt zu erschließen. Unter Navigation versteht man die geleitete Bewegung von einem Knoten in einem Navigationsgraphen zu einem benachbarten Knoten, wobei jeder Knoten die Beschreibung eines Datenobjekts oder Metadatum repräsentiert. Eine Schlagwortsammlung ist beispielsweise eine übliche Struktur, die zu Navigationszwecken herangezogen wird.

2.4 Benutzergerechte Aufbereitung von gefundenen Inhalten

Ein Nutzer wird bei seiner Suche nach relevanten Daten zunächst im Metadatenbestand suchen bzw. navigieren, um später Details zu den gefundenen Informationen abzurufen. Die Rückmeldungen, die das System dem Benutzer als erste Antwort seiner Suche bzw. Navigation liefert, sind jeweils adäquat an die Situation anzupassen. Beispielsweise werden Informationen über die räumliche Ausdehnung eines Datenbestandes in Karten, eine Aufstellung verschiedener Meßmethoden eher in Tabellen mit kurzen Erklärungen geliefert.

Auch die angeforderten Rohdaten, die auf eine Anfrage des Benutzers passen, sollen in möglichst geeigneter Weise präsentiert werden. Bevor Rohdaten über das Netzwerk verschickt werden, soll dem Nutzer ausreichend Information darüber zur Verfügung stehen, ob diese Daten tatsächlich relevant sind und welchen Umfang die Datenübertragung haben wird. Es ist davon auszugehen, daß Rohdaten durch den Benutzer rechnerunterstützt weiterverarbeitet werden, wofür in den meisten Fällen Standardsoftware zum Einsatz kommt. Deshalb ist eine Ausgabe in Standardformaten vorzusehen. Umfangreiche Dokumente können ausschnittsweise angefordert werden. Eine Ausgabe in der ursprünglichen Rohdatenform soll nur auf besondere Anfrage hin erfolgen, beispielsweise bei Spezialformaten für Satellitenbilddaten.

2.5 Beachtung von Zugriffsbeschränkungen

Nicht alle verfügbaren Daten stehen kostenlos oder öffentlich zur Verfügung. Sicherheitsaspekte und juristische Gründe (Datenschutz, Urheberrechte) können dafür sprechen, eine Kontrolle der Zugriffe einzurichten. Grundsätzlich werden im Kontext des Projektes TIDE Metadaten frei verfügbar sein, d.h. Informationen darüber ob und wo relevante Daten vorliegen unterliegen keinerlei Zugriffsbeschränkungen.

Für den sicheren Zugang zu Rohdaten ist ein Authentisierungsverfahren vorgesehen. Ein Endbenutzer erhält einen elektronischen Zugangsschlüssel und weist sich zu Beginn der Systemnutzung damit aus. Im weiteren Verlauf der Nutzung sind dann keine weiteren Berechtigungsnachweise vom Endbenutzer einzugeben. Es ist ein Rollenkonzept vorgesehen, das einen Benutzer verschiedenen Nutzergruppen zuordnet. Die Verwaltung konkreter Nutzungsrechte für einzelne Nutzergruppen obliegt dann den einzelnen Servern im Gesamtsystem. Falls eine verschlüsselte Übertragung von Daten notwendig ist, so kann dies auf Basis einer sicheren Transportschicht (SSL) geschehen. Für jeden Datenbestand ist die Möglichkeit vorgesehen, Zugriffe durch nicht autorisierte Personen zu unterbinden. Weiterhin werden alle Zugriffe protokolliert, um Versuche mißbräuchlicher Nutzung zu erkennen und entsprechend reagieren zu können.

3 Konzeption und Architektur eines Zugriffsystems

In diesem Abschnitt wird eine CORBA-basierte, dreistufige Client-Server-Architektur vorgestellt, die den transparenten Zugriff auf verschiedene Datensammlungen ermöglichen soll. Als Benutzungsschnittstelle fungiert in der ersten Stufe ein WWW-Browser, über den verschiedene Applets von einem WWW-Server heruntergeladen werden. Diese Applets enthalten graphische Benutzungs-elemente, die die Kommunikation mit den anderen Komponenten des TIDE-Systems anstoßen. Sie beinhalten außerdem Kopplungsmethoden zu den Middlewarekomponenten der zweiten Stufe. Die dritte Stufe besteht aus den proprietären Datenhaltungssystemen wie z.B. Datenbanken oder Dateisysteme.

Die Kommunikation wird über CORBA-konforme Komponenten abgewickelt. CORBA ist ein Standard für Architektur und Schnittstellen in verteilten Systemen. Wir stützen uns auf die Version 2.0, für die uns die Implementation eines Object Request Brokers (ORB) zur Verfügung steht (VisiBroker). Der ORB gewährleistet das Verschicken von Nachrichten und das Umsetzen von Typinformationen. Er kann auch Dienste wie Naming, Event, Security u.a. übernehmen (Abb. 1).

Das TIDE-System sieht verschiedene Arten von Komponenten der mittleren Stufe vor, die auf der Basis von CORBA-Objekten durch Server bereitgestellt werden: Es stehen Anfrage-, Daten-, Thesaurus-, Karten- und Konvertierungsserver zur Verfügung. Die Benutzungsschnittstellen der Applets sind im nächsten Abschnitt

beschrieben. Im folgenden sollen kurz die einzelnen Server und ihre Dienste dargestellt werden.

Ein Anfrageserver, der zur Lastverteilung auch in mehreren Instanzen vorliegen kann, vermittelt zwischen Anfragen, die im Browser zusammengestellt wurden, und den Objekten eines Datenservers, der von einer datenhaltenden Stelle implementiert wird. Eine Anfrage liefert zunächst eine Übersicht aller passenden Datenobjekte. Sie wird als Liste mit kurzen beschreibenden Informationen (Objektidentifikation, Abstract, mögliche Formate, Verweise auf weitere Informationsquellen) im Browser dargestellt.

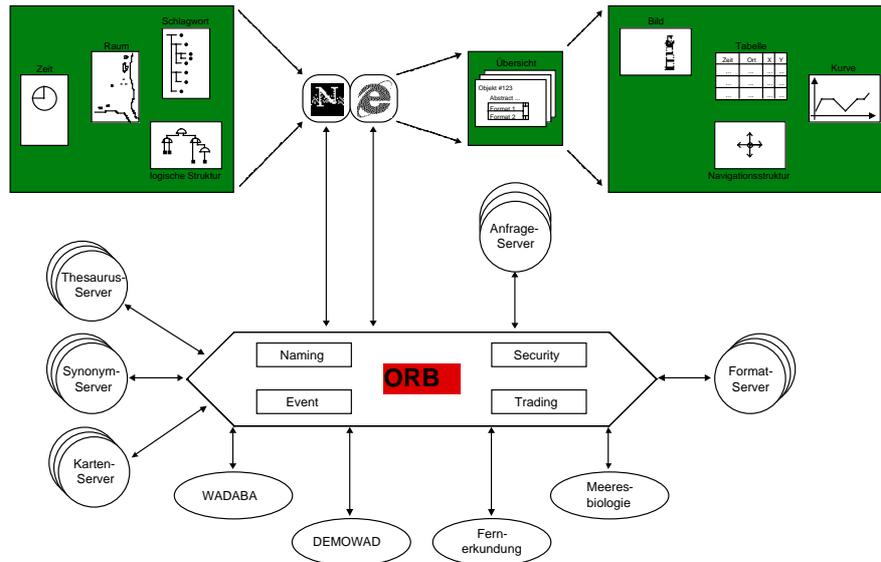


Abb. 1: Architektur des Gesamtsystems

Indem der proprietäre Zugang zum Datenbestand einer datenhaltenden Stelle in eine spezielle CORBA-Kapsel eingeschlossen wird, steht der Datenbestand als Menge von Objekten eines Datenservers zur Verfügung. Diese Objekte können mit dem übrigen TIDE-System über die vorgesehenen Schnittstellen kommunizieren. Ein Datenserver beantwortet eingehende Anfragen bzgl. seines Datenbestandes. Er ermöglicht damit sowohl die Suche nach vorgegebenen Kriterien als auch ein Navigieren durch den Datenbestand und liefert gefundene Metadaten oder Rohdaten. Zunächst kommen Datenserver für die folgenden Datenbestände in Betracht:

- Wattenmeerdatenbank WADABA (Krasemann et al. 93)

- Wattenmeermonitoring-Datenbank DEMOWAD (Pommerencke et al 97)
- die Elbe-Datenbank DROPS (Zabanski 96)
- Satellitenbilder der GKSS-Abteilung Fernerkundung²

Über einen Thesaurus-Server kann jede datenhaltende Stelle eigene Schlagwort-sammlungen zur Verfügung stellen. Damit können Datenobjekte korrekt verschlag-wortet und gezielte Suchanfragen unterstützt werden. Weiterhin besteht die Möglich-keit, diese Sammlungen als Grundlage für eine Navigationsstruktur zu verwenden, um die Menge der Datenobjekte zu durchstöbern. Im Gesamtsystem werden mehrere Schlagwortsammlungen nebeneinander vorhanden sein, falls beispielsweise mehrere Sprachen unterstützt werden oder Betreiber Autonomie bzgl. ihrer Schlagwortsamm-lungen beanspruchen. Im Einvernehmen der einzelnen Betreiber dieser Schlagwort-sammlungen können jedoch schrittweise Beziehungen zwischen Einträgen aufgebaut und manuell in das Gesamtsystem integriert werden. Ein Synonymserver verwaltet diese Beziehungen, der neben Verwandtschaftsbeziehungen auch Übersetzungen anbieten könnte.

Sind räumliche Informationen ein- oder auszugeben, so stehen Server bereit, die das notwendige Kartenmaterial zur Verfügung stellen. Dabei ist nicht an die Funktionalität eines geographischen Informationssystems gedacht, das aus eingele-senen Einzelinformationen automatisch Karten erzeugen und Abfragen beantworten kann. Es sollen zunächst lediglich Hintergrundbilder zur Ein- oder Ausgabe von Koordinaten unterstützt werden. Im einfachsten Fall sind dazu Satellitenfotos mit hoher Auflösung vorhanden. Diese Bilder können dann gestückelt oder in geringer Auflösung bereitgehalten werden, um den notwendigen Datenverkehr zur Übertra-gung eines gewünschten Ausschnittes moderat zu halten.

Jedes per Suchanfrage oder Navigation gefundene Datenobjekt soll letztendlich in den Browser geladen, dort angezeigt und ggf. direkt auf das lokale Dateisystem beim Nutzer kopiert werden. Da auf der Browserseite evtl. andere Formate unterstützt werden als auf der Seite eines Datenservers, sind Funktionen zum automatischen Konvertieren von Datenobjekten wünschenswert (z.B. Konvertierung von Satelliten-rohdaten in Standardformate wie GIF, JPEG o.ä.). Ein Formatserver liefert diese Funktionalität für ein oder mehrere Quell- und Zielformate.

Durch die Kombination von Daten- und Formatserver ist es möglich, den Export von Daten oder Metadaten durchzuführen, um sie auch für andere Umweltinforma-tionssysteme verfügbar zu machen, die die CORBA-Schnittstellen nicht direkt nutzen können. Ein Datenserver stellt dann die Information über den gesamten Datenbestand oder über einen spezifischen Ausschnitt zusammen und liefert ein SGML-Dokument ab. Aus diesem Dokument kann der Formatserver die speziell für das gewünschte Format benötigten Informationen extrahieren und in das Zieldoku-ment aufnehmen.

² <http://w3g.gkss.de/G/GFE.html>

4 Implementationsansätze

Die bisherigen Implementationsarbeiten beschäftigten sich mit einer einfachen Schnittstelle zu der relationalen Wattenmeerdatenbank WADABA im Forschungszentrum GKSS. Daten der WADABA sind in Projekten organisiert und können mit Hilfe eines proprietären Schlagwortkatalogs (Menübaum) aufgefunden werden.

Es wurde eine Navigationsstruktur in Form von statischen HTML-Seiten erstellt, die die Organisationsstruktur der Wattenmeerdatenbank abbildet. Dieser Satz von HTML-Seiten, die Projektbeschreibungen enthalten, wird durch eine Suchmaschine recherchierbar gemacht. Der Zugriff auf Rohdaten in der Datenbank erfolgt durch CGI-Skripte. Es können einzelne Tabellen der Projekte komplett oder in Teilen ausgelesen und als HTML-Seite formatiert werden.

Dieser erste Prototyp wurde zügig realisiert, um Rückmeldungen von Benutzern über den erreichbaren Datenbestand und die gewünschten Zugriffsarten einzuholen. Es zeigte sich, daß dieses erste System einfach zu benutzen ist und Inhalte übersichtlich darstellt. Die Recherche nach Datenbeständen bietet jedoch nur einen geringen Funktionsumfang, da lediglich die beschreibenden Texte der Projekte erschlossen werden. Es fehlen Suchmöglichkeiten, die einen Raum- und Zeitbezug quantifizieren sowie Inhalte von Rohdaten berücksichtigen. Auch die Anbindung von Datenbeständen anderer Organisationen ist mit dieser Systemversion nicht möglich.

Deshalb werden zur Zeit Arbeiten mit Hilfe einer aktuellen CORBA 2.0-Implementierung durchgeführt. Dabei sind die technischen Möglichkeiten und Restriktionen des Einsatzes von Java-Applets als Realisierungsgrundlage einer neuen Benutzungsschnittstelle auszuloten. Im folgenden sollen die Aspekte der Benutzungsschnittstelle und der Datenübertragung genauer erläutert werden.

4.1 Graphische Benutzungsoberfläche

Es gibt zwei Einstiegspunkte für einen Endbenutzer: Entweder läßt er eine Suchanfrage auf ausgewählten Datenservern durchführen, oder er selektiert einen Datenserver und durchstöbert dessen Datenbestand anhand von Navigationsstrukturen, die ihm der Datenserver anbietet.

Zur Eingabe einer Suchanfrage steht ein Editor zur Verfügung, der die logische Struktur der Anfrage aufbaut, die Attributwerte der Suchkriterien ausfüllt und Prädikate auswählt, die für die Attributwerte zu überprüfen sind (Abb. 2). Für jedes Suchkriterium ist ein eigenes Ausfüllmuster in Form eines Dialogs vorgesehen. Wählt der Benutzer also ein Suchkriterium aus, so erscheint zunächst das entsprechende Dialogfenster. Ein typisches Beispiel für solch einen Dialog ist die Angabe von geographischen Koordinaten über eine Karte (Abb. 3). Sind die Angaben für Prädikat und Attribute eingetragen, so erscheint ein benannter Verweis auf diesen

Teil der Anfrage im Editor. Alle Einzelteile einer Anfrage lassen sich mit Hilfe von logischen Operatoren zu einer komplexen Suchanfrage zusammenstellen.

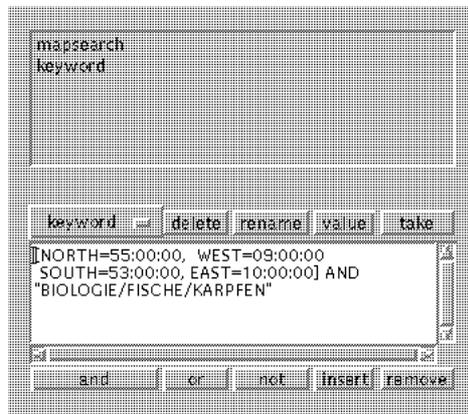


Abb. 2: Applet für die Eingabe von Anfragen

Das Ergebnis einer Suchanfrage ist eine Liste mit Kurzbeschreibungen der gefundenen Datenobjekte (Abb. 4). Für jedes Objekt kann in einem weiteren Schritt eine erweiterte Beschreibung angefordert werden (Abb. 5). Beim Navigieren erscheint ebenfalls für jeden Knoten im Navigationsgraphen eine Kurzbeschreibung, von der aus die erweiterte Beschreibung erreicht werden kann. In der erweiterten Beschreibung sind die Attribute durch entsprechende Werkzeuge einzeln anzeigbar. Typische Beispiele hierfür sind wieder die Angabe geographischer Koordinaten sowie die Stellung eines Objekts innerhalb einer Navigationsstruktur.

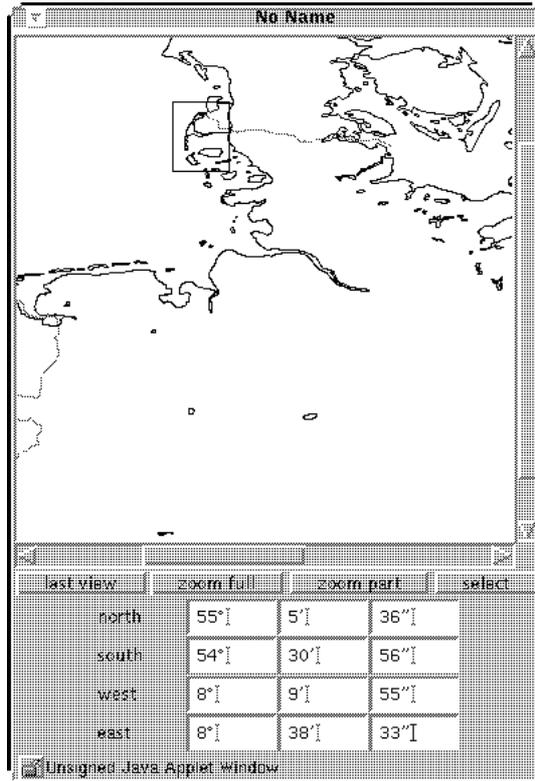


Abb. 3: Zoom-Applet für die Eingabe geographischer Koordinaten

Result Overview for Query:
 ((time > 1990) AND
 (position intersects [1234567890; 007; 0815; 4711]))

Name	Mime Type	Size	Matched Parameters
Fish	text/html	1024	time = 1997 position.north = 53-06-00 position.south = 53-06-00 position.west = 08-03-00
Seal	image/jpeg	1234567890	position.south = 53-06-00 position.west = 08-03-00 position.east = 09-00-00
Breeding Birds	sound/av	4711	time = 1997 position.north = 53-06-00 position.south = 53-06-00 position.west = 08-03-00

Show Details

Abb. 4: Anfrageergebnisse im Überblick

No Name		
Result Details for Query:		
(time > 1990) AND (position intersects [1234567890; 007; 0815; 4711])		
Name	Value	View
Name	Seal	View
Abstract	Full descriptional text	View
Contact	Ralf Kriebisch University of Hamburg	View
Time Span	01.01.1900 – 31.12.1999	Showtime
Geographical Coordinates	NORTH= 55:01:02 SOUTH= 53:03:04	Zoom
Matched	time = 1997 position.north = 53-05-00	View
Mime-Type	image/jpeg	Earl
Size	1234567890	Grey
Origin	loadtp://dataserver.sub.domain/path/file.ext	Download

Abb. 5: Anfrageergebnisse im Detail

Zum Navigieren im Datenbestand eines Datenservers kommt ein spezieller Graph-Browser zum Einsatz, der innerhalb des Java-Applets im WWW-Browser abläuft (Abb. 6). Die Navigationsstruktur des Datenservers ist ein Graph mit einem festgelegten Einstiegsknoten. Ein Knoten enthält eine Datenbeschreibung, die Kanten beschreiben verschiedene Arten von Beziehungen zu anderen Knoten. Der Graph-Browser stellt den jeweils aktuellen Knoten dar und ermöglicht es, einen benachbarten Knoten einzustellen. Um die Nachbarschaftsbeziehungen darzustellen, sind vielfältige Ansätze möglich, z.B. aufwendige 2D- oder 3D-Darstellungen. Zunächst wird jedoch eine sehr einfache Lösung auf Standard GUI-Elementen wie Pop-Up-Menüs und Auswahllisten erstellt.

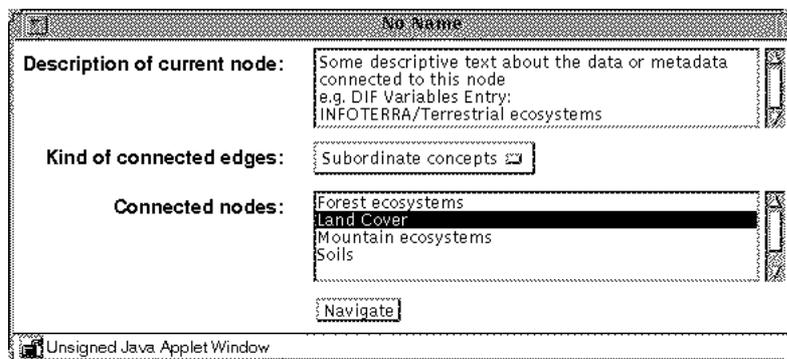


Abb. 6: Graph-Browser für Schlagwortbäume

Im Prinzip könnte auch der WWW-Browser als Navigationsinstrument benutzt werden. Dabei wäre der Zugriff auf die Knoten des Navigationsgraphen jedoch beschränkt auf URLs, und es müßte eine Zwischenlösung wie z.B. CGI oder Servlets gewählt werden, um den Datenserver über CORBA anzubinden. Dieses Vorgehen halten wir nicht für sinnvoll, da eine Codierung der Zustandsinformation des Clients vorgenommen werden muß und sich CGI-Skripte oder Servlets leicht als Flaschenhals in einer Gesamtarchitektur erweisen.

4.2 Übertragung von Rohdaten auf das lokale Client-System

Da Rohdaten häufig lokal weiterverarbeitet werden, muß ein Mechanismus vorgesehen sein, um einzelne Datenobjekte oder aggregierte Datensätze beim Client als Datei zu speichern. Aus einem Java-Applet heraus ist jedoch das Schreiben auf das lokale Dateisystem nicht ohne weiteres möglich. Bei einigen Browsern lassen sich spezielle Sicherheitsrestriktionen der Java-Umgebung zwar explizit lockern, dies ist jedoch bis Ende des TIDE-Projekts nicht von allen eingesetzten WWW-Browsern zu erwarten. Deshalb wird eine Zwischenlösung auf der Basis von dynamisch erzeugten URLs eingeführt (Abb. 7).

Fragt der Client ein Datenobjekt als lokale Kopie nach, so wendet er sich an den Datenserver, der in der benutzten CORBA-Implementation über einen Stellvertreter (Gatekeeper) angesprochen wird. Der Datenserver stellt dieses Objekt bereit und beauftragt einen Formatserver, aus dem Datenobjekt ein Dokument zu erstellen, das als Datei weiterverarbeitet werden kann. Der Formatserver speichert das Dokument, erzeugt einen eindeutigen URL für dieses Dokument und gibt diesen URL zurück. Der Client kann nun diesen URL aufrufen und der Endbenutzer erhält die Möglichkeit, den Inhalt des Dokuments per Menübefehl des WWW-Browsers im lokalen Dateisystem abzulegen. Hat der Formatserver keinen direkten Zugang zu einem Dateisystem, von dem aus gängige Internet-Server (http- oder ftp-Server) die abgelegten Dokumente weiterleiten können, so ist ein weiterer Zwischenschritt über CGI oder Servlets notwendig, um das Dokument als URL zugänglich zu machen.

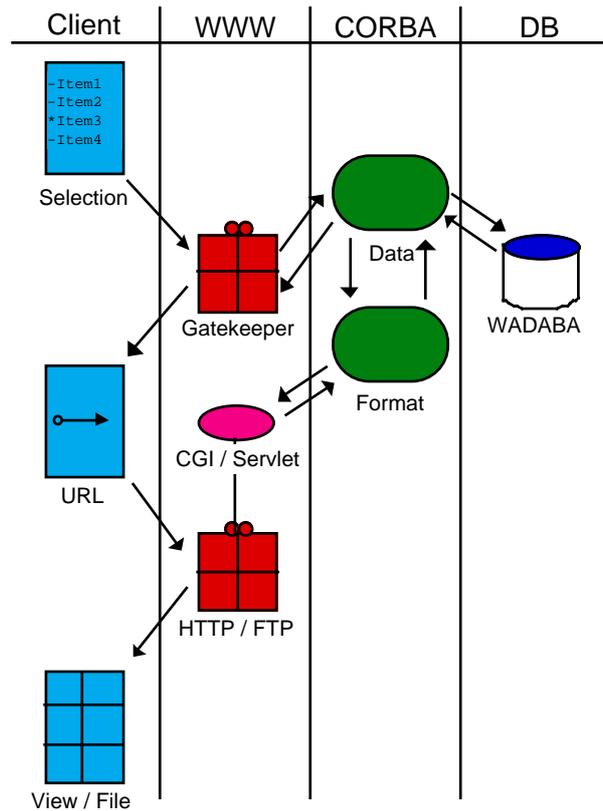


Abb. 7: Zwischenschritte der Rohdatenübertragung

5. Ausblick

Im weiteren Projektverlauf steht zunächst die Implementation wichtiger konzipierter Dienste im Vordergrund. Exemplarisch werden die Wattenmeerdatenbank WADABA und die Elbe-Datenbank DROPS vollständig als Datenserver realisiert. Sie unterstützen dann auch die Auslieferung von Rohdaten an den Endbenutzer. Ein Kartenserver wird Bildmaterial aus Satellitenfotos in verschiedenen Auflösungen liefern.

Im Rahmen der Implementierung eines Editors zur Erstellung von Anfragen ist eine eigene Sprache zu formalisieren, die die aussagenlogische Struktur dieser Anfragen beschreibt. Damit können Anfragen als Texte an Anfrageserver verschickt, dort analysiert und durch Datenserver systematisch abgearbeitet werden.

Nach der Definition und Einführung von Nutzer- und Datenbestandsprofilen (z.B. Themen oder Präferenzen zur Datenaufbereitung) könnten aus dem Bereich

'electronic commerce' bekannte Trading-Verfahren genutzt werden. Diese Verfahren versuchen, Nutzer- und Datenbestandsprofile passend zuzuordnen und ermöglichen so das Auffinden gesuchter Datenobjekte.

Die bisherigen Arbeiten erstrecken sich ausschließlich auf die Seite der Datenausgabe bereits erfaßter Information. Auch zur Dateneingabe ist eine komfortable GUI-Unterstützung sinnvoll. Es ist noch zu untersuchen, ob die Eingabe von Daten ebenfalls über neue Dienste des Gesamtsystems geleistet werden könnte, um von den Grunddiensten zu profitieren, die eine ausgefeilte CORBA-Implementierung anbieten kann.

Abschließend bleibt festzuhalten, daß Umweltinformationssysteme bei der Akquisition und der Aufbereitung von Daten nur eine unterstützende Funktion haben. Die Bewertung dieser Daten und die Reaktion auf deren Inhalt obliegt jedoch weiterhin den Endbenutzern. Wünschenswert ist, daß von einer überzeugenden Darstellung relevanter Information zur richtigen Zeit Entscheidungen ausgehen, die der Umwelt und damit der in ihr lebenden Menschen zugute kommen.

Literatur

- B. Gehlsen et al. (1998): Das Projekt TIDE - Werkzeuge für eine einheitliche Sicht auf verteilte, heterogene Umweltdaten; in: J. Hoppe et al. [Hrsg.]: *Vernetzte Umweltinformation*, Metropolis-Verlag, Marburg.
- H. L. Krasemann et al. (1993): Nutzerführungssysteme für das Wattenmeerinformationssystem WATiS, in: A. Jaeschke et al. [Hrsg.]: *Informatik für den Umweltschutz - 7.* Symposium, Ulm, - Proceedings. Informatik-aktuell, Springer-Verlag, Berlin, 1993.
- W. Lamersdorf (1994): *Datenbanken in verteilten Systemen*, Vieweg; Braunschweig, Wiesbaden.
- R. Orfali, D. Harkey (1997): *Client/Server Programming with JAVA and CORBA*, John Wiley & Sons; New York et al..
- B. Page, L.M. Hilty (Hrsg.) (1995): *Umweltinformatik*, Oldenbourg; München; 2. akt. u. erw. Auflage.
- M. Pommerencke et al. (1997): Networking and Harmonisation of International Monitoring Data, in: W. Geiger et al (Hrsg.): *Umweltinformatik '97*, Metropolis-Verlag, Marburg.
- M. Rech (1998): "Rohstoff" Umweltdaten - mit Informationssystemen zum Verbraucher, GeoSpektrum 3/98, Zeitschrift der Alfred-Wegener-Stiftung.
- S. Zabanski (Hrsg.) (1996): *DROPS Elbe 1.0 für Windows™*, Die Elbe-Datenbank des Sonderforschungsbereichs 327 auf CD-ROM, Institut für Meereskunde, Zentrum für Meeres- und Klimaforschung, Universität Hamburg.