

Datenbankkonzepte und -systeme in Umweltinformationssystemen (UIS)

Franz Josef Radermacher und Wolf-Fritz Riekert, FAW Ulm

In Anbetracht der Bedrohungen der heutigen Welt, insbesondere durch Überbevölkerung, wachsenden globalen Konsum und zunehmend gefährliche Arten von Abfall, sind verlässliche und umfassende Informationen über den Zustand unserer Umwelt von unschätzbarem Wert. Die Bereitstellung und die Verfügbarmachung derartiger Informationen machen in zunehmendem Maße eine informationstechnische Unterstützung durch Umweltinformationssysteme erforderlich. Umweltdatenbanken, als Basiskomponenten solcher Systeme, dienen zur Strukturierung, Verwaltung und Ablage der Umweltinformationen für die informationstechnische Verarbeitung. Dabei werden, über die Eigenschaften von herkömmlichen Datenbanken hinaus, zusätzliche Datenkonzepte und Verwaltungsmechanismen erforderlich, die sich insbesondere auf die erforderliche Metainformation, die räumlich-zeitlichen Aspekte und die Abfragesprache beziehen.

Database Concepts and Systems in Environmental Information Systems

Information about the environment is increasingly being made available by means of software systems known as Environmental Information Systems. Environmental Databases, as base components of these systems, serve for the storage and management of environmental information. Beyond the system features of traditional databases, environmental databases require new data concepts and management methods, particularly with regard to the necessary meta-information, the spatio-temporal aspects, and the query language.

1. Daten und Information

In der öffentlichen Diskussion wird häufig Wert auf die Unterscheidung zwischen *Daten* und *Information* gelegt [14]. Gerade im Umweltbereich ist von nutzlosen „Datenfriedhöfen“ die Rede, wobei gleichzeitig eine riesige „Informationslücke“ beklagt wird. Dennoch ist es schwierig, den Unterschied zwischen Daten und Information mit formalen Methoden zu fassen. Im allgemeinen sind mit Daten recht einfache und elementare Aspekte verbunden. Man denke nur an Meßwerte von Schadstoffkonzentrationen oder an Bevölkerungszahlen in bestimmten Gebieten. In der Regel gibt es wohldefinierte Methoden, mit denen derart elementare Werte verifiziert werden können, gleichwohl, ob es sich um Schadstoffkonzentrationen oder Bevölkerungszahlen handelt. Mit Daten ist also die Vorstellung verbunden, daß es sich hierbei um etwas Festes, Klares und Unanzweifelbares handelt, etwas, das auf eine standardisierte Weise ermittelt oder abgeleitet werden kann. Im Gegensatz hierzu erscheint Information als etwas, das einer höheren Ebene angehört, etwas, das nicht leicht gemessen oder ermittelt werden kann, sondern etwas, das auf eine schwierige und nicht unmittelbar offenkundige Weise im Hinblick auf besondere Fragen aus vielen Einzeldaten abgeleitet werden muß. Entscheidungsträger sind in den meisten Fällen an Information in dem beschriebenen Sinne interessiert und nicht an einfachen (Roh-)Daten. So sind sie z.B. an Information über die Luftqualität in einer bestimmten Gegend interessiert oder an Information über die Entwicklung der Sozialstruktur eines bestimmten Wohngebiets im Verlauf der Zeit. Fragen der Lebensqualität, der Sozialstruktur, der Landesentwicklung usw. gehören zu den Informationsbedürfnissen, die vergleichsweise schwierig abzudecken sind und stehen in einem engen Zusammenhang mit dem Anwendungsgebiet des jeweiligen betrachteten Umweltinformationssystems.

2. Metainformation

Eine Umweltdatenbank, als Teil eines Umweltinformationssystems, hat idealerweise die Aufgabe, alle Daten bereitzuhalten, aus denen – mit Hilfe der in dem Umweltinformationssystem verfügbaren Methoden – die interessierenden Umweltinformationen abgeleitet werden können. Diese Idealvorstellung ist freilich in den meisten heute realisierten Umweltinformationssystemen noch nicht Stand der Technik. In vielen Fällen werden noch nicht einmal Datenbanken benutzt. Es wird vielmehr auf einfache Datensammlungen, z.B. Magnetbänder, zurückgegriffen, die die Massendaten (z.B. Bilddaten oder Meßdaten) enthalten. Alle weiteren zur Interpretation dieser Daten erforderlichen Informationen muß der Sachbearbeiter beisteuern oder sie stecken implizit in den Auswertungsprogrammen. Ein solches Magnetband ist Jahre später, etwa nach einem Systemwechsel oder dem Weggang des Sachbearbeiters, praktisch wertlos, weil dann die Zusatzinformation fehlt, die erforderlich ist, die Daten sachgerecht zu interpretieren. Aus heutiger Sicht sollte deshalb die gesamte zur Interpretation erforderliche Information in einer Umweltdatenbank bereitgehalten werden. Neben den „eigentlichen“ Umweltdaten sind dies u.a. Vermerke über Ortsbezug, Zeitbezug, Genauigkeit, Bearbeitungsstand, aber auch über die verwendeten Datenstrukturen und Datenformate oder über spezielle Codierungstechniken. Diese Art von zusätzlicher Information wird oft auch als sogenannte Metainformation bezeichnet. Diese Bezeichnung sollte nicht dahingehend mißverstanden werden, daß es sich dabei um eine neue Art von Information handelt. Bei der sogenannten Metainformation handelt es sich um gewöhnliche Informationen, die auch auf normale Weise in einer Datenbank abgelegt werden kann, nur bezieht sie sich auf ein anderes Themengebiet, d.h., nicht mehr auf die konkrete Umwelthematik, sondern auf die Qualität und die Struktur der über diese Umwelthanwendung vorhandenen Daten. Es ist nach heutigem Wissensstand unbestritten, daß auch solche Metainformationen in einer Datenbank verfügbar gemacht werden sollten [19].

3. Umweltdatenbanken

Technisch gesehen handelt es sich bei einer Umweltdatenbank um eine Datenbank im herkömmlichen Sinne. Das heißt insbesondere, daß eine Umweltdatenbank alle in der Fachliteratur beschriebenen Eigenschaften aufweisen sollte, die ein Datenbanksystem üblicherweise kennzeichnen. Zu diesen Eigenschaften zählen Konzepte wie persistente Datenhaltung, die Verfügbarkeit einer Datenbanksprache, die Unterstützung des Mehrbenutzerbetriebs, das Vorhandensein eines Transaktionskonzepts sowie Mechanismen zur Datensicherheit. Die Natur der Umweltdaten sowie die Besonderheiten von Umwelthanwendungen erfordern allerdings – über die Eigenschaften von herkömmlichen Datenbanksystemen hinaus – zusätzliche Datenkonzepte und Verwaltungsmechanismen [4]. Solche Konzepte und Mechanismen waren auch Gegenstand einer Reihe von Forschungsprojekten, die am Forschungsinstitut für anwendungsorientierte Wissensverarbeitung (FAW) an der Universität Ulm mit Unterstützung der Stifter des FAW sowie in Zusammenarbeit mit dem baden-württembergischen Umweltministerium im Kontext des Umweltinformationssystems (UIS) Baden-Württemberg durchgeführt wurden und auf die in den folgenden Abschnitten noch besonders eingegangen wird.

4. Umweltdaten

Umweltdaten haben viele Eigenschaften, die sie von herkömmlichen Daten, etwa in kaufmännischen Anwendungen, unterscheiden. So sind Umweltdaten häufig raumbezogen, d.h. ihre Gültigkeit bezieht sich auf bestimmte geographische Orte. Die Daten sind zumeist auch zeitbe-

zogen, d.h. sie beziehen sich auf einen Zeitpunkt oder Zeitraum. Eine ganz besondere Kategorie von Daten bilden Meßdaten, die für diskrete Zeitpunkte und Orte erhoben werden und die für eine Modellierung kontinuierlicher Verläufe erst noch interpoliert werden müssen; ein Beispiel hierfür sind die zeitlich und räumlich verteilten Meßwerte, die die Luftmeßnetze des Bundes und der Länder liefern.

Ein Spezialfall von Meßdaten sind Bilddaten, wie sie z.B. von Satellitensensoren mit Mitteln der Fernerkundung erhoben werden. Bilddaten weisen gegenüber anderen Arten von Daten eine Reihe von Besonderheiten auf. Zum einen gehört zu diesen Bilddaten eine große Anzahl von beschreibenden Daten, etwa zum verwendeten Sensor, zum Ort und zur Zeit der Aufnahme, zur Projektionsart usw. Zum anderen handelt es sich bei den Bilddaten vom Umfang her um sehr große Datenmengen, die in den Bereich der Tera-Bytes gehen. Dieser Umfang liegt zwei Größenordnungen oberhalb der Menge von Daten, die heute im Bereich der Finanztransaktionen bzw. in den großen internationalen Buchungssystemen anfallen. Dies erfordert neue Speichermodelle beim Datenbankdesign, insbesondere die Verwendung von sogenannten Tertiärspeichern, z.B. auf der Basis vollautomatischer CD-Plattenarchive.

Umweltdaten sind häufig sehr komplex, d.h. sie beschreiben komplexe Umweltobjekte. Solche Umweltobjekte sind in der Vorstellung des Menschen a priori häufig gar nicht vorhanden, im Gegensatz zu Objekten aus der technischen Welt müssen sie oft erst noch gebildet werden, und zwar sowohl in begrifflicher wie in informationstechnischer Hinsicht. Beispiele hierfür sind die Abgrenzung von Umweltobjekten auf Satellitenbildern oder die Identifizierung von Homogenitätsbereichen bei der Grundwassermodellierung. Eine weitere Schwierigkeit besteht darin, daß Umweltdaten häufig mit Unsicherheit behaftet sind. Deshalb muß es möglich sein, mit den eigentlichen Daten Toleranzen, Wahrscheinlichkeiten oder Konfidenzfaktoren abzuspeichern. Schließlich sind Umweltdaten häufig verteilt auf heterogene Datenbanksysteme. Dies erfordert Integrationsleistungen, wie sie beispielsweise in den FAW-Projekten WINHEDA (Wissensbasierter Zugriff auf verteilte heterogene Wissensquellen [8]) und INTEGRAL (Integration heterogener UIS-Komponenten [21]) erbracht werden.

5. Räumliche Aspekte von Umweltdaten

Eine wichtige Aufgabe von Umweltinformationssystemen ist die Verarbeitung raumbezogener Umweltdaten. Hierfür werden typischerweise sogenannte Geoinformationssysteme eingesetzt. Geoinformationssysteme zeichnen sich dadurch aus, daß sie auf äußerst komplexen Geometrie- und Sachdatenstrukturen operieren. Moderne am Markt erhältliche Geoinformationssysteme nutzen relationale Datenbanksysteme zur integrierten Ablage dieser Geometrie- und Sachdaten. Darüber hinaus ist die Nutzung neuer, sogenannter postrelationaler Datenbanksysteme [6] Gegenstand der Forschung; insbesondere die seit einigen Jahren auf dem Markt verfügbaren objektorientierten Datenbanksysteme versprechen zusätzliche Vorteile für die Repräsentation von raumbezogener Umweltinformation [9].

Am FAW wird im Rahmen des u.a. von der Siemens Nixdorf Informationssysteme AG und vom Umweltministerium Baden-Württemberg in Auftrag gegebenen Projekts GODOT (Geodatenhaltung mit objektorientierten Techniken [20]) der Prototyp eines Geoinformationssystems entwickelt, das direkt auf einem dieser objektorientierten Datenbanksysteme aufsetzt. In der Terminologie von GODOT lassen sich drei Kategorien von Daten unterscheiden, die einem Geoinformationssystem zugrundeliegen, nämlich thematische Daten, geometrisch-topologische Daten und Präsentationsdaten. Alle drei Kategorien von Daten lassen sich sehr gut in objektorientierter Form darstellen, nämlich in Form sogenannter thematischer Objekte, geometrischer Objekte und Präsentationsobjekte.

Thematische Objekte sind einfache oder komplexe (d.h. mit Hilfe anderer Objekte definierte) Objekte, in deren Attributen geographische und umweltbezogene Sachinformation repräsentiert wird. Thematische Objekte können unterschiedliche Arten von Attributen besitzen:

- Attribute, die elementare Werte besitzen (z.B. Zeichenketten oder numerische Werte);
- Attribute, die komplexe Werte besitzen (realisiert durch Unterobjekte, die in das thematische Objekt eingebettet sind);
- Attribute, die Referenzen auf einzelne oder mehrere andere eigenständige Datenobjekte (insbesondere andere thematische Objekte) als Werte besitzen.

Die wichtigste Art von thematischen Objekten sind die sogenannten *Geoobjekte*. Gegenüber einem gewöhnlichen thematischen Objekt (z.B. einer Tierart) zeichnet sich ein Geoobjekt dadurch aus, daß es neben den oben genannten Arten von Attributen noch eine sogenannte Geometrie besitzt, durch die eine geographische Lage und Ausdehnung repräsentiert wird.

Die Geometrie eines Geoobjekts läßt sich im allgemeinsten Fall als Menge sogenannter *geometrischer Objekte* repräsentieren. In GODOT, ähnlich wie auch in den meisten anderen Geoinformationssystemen, gibt es drei Arten von elementaren geometrischen Objekten, nämlich flächenhafte, linienhafte und punktförmige geometrische Objekte; eine vierte Art von geometrischen Objekten, die räumlichen Körper, wird für die Erweiterung des Systems auf dreidimensionale Anwendungen (z.B. in der Geologie) gebraucht. Topologische Beziehungen zwischen solchen geometrischen Objekten lassen sich durch bidirektionale Verzeigerungen repräsentieren. So darstellbar ist dann z.B. die Tatsache, daß ein punktförmiges geometrisches Objekt Anfangs- oder Endpunkt eines linienhaften geometrischen Objekts ist, oder, daß ein linienhaftes geometrisches Objekt einen Teil des Umrisses eines flächenhaften geometrischen Objekts darstellt.

Präsentationsobjekte dienen zur graphischen Darstellung von Sachobjekten, insbesondere von Geoobjekten. Diese graphische Darstellung könnte grundsätzlich von ganz allgemeiner Art sein, z.B. kann eine Business-Graphik, eine Tabelle oder eine Videosequenz, die ein Sachobjekt darstellt, durch ein Präsentationsobjekt repräsentiert sein. Die wichtigste Art von Präsentationsobjekten sind die kartographischen Objekte. Diese verkörpern alle beschreibenden Daten, die erforderlich sind, um Geoobjekte in kartographischer Form darzustellen. Diese in GODOT betonte Trennung zwischen geographischer und kartographischer Information entspricht auch der im Amtlichen Topographisch-Kartographischen Informationssystem ATKIS [1] vorgenommenen Unterscheidung eines digitalen Landschaftsmodells (DLM) und eines digitalen kartographischen Modells (DKM). Kartographische Objekte sind durch graphische Attribute (wie z.B. Farbe, Strichstärke, Schriftarten) gekennzeichnet. Diese Informationen haben mit den Objekten der realen Welt, die durch Geoobjekte repräsentiert werden, nicht unmittelbar etwas zu tun. Es ist infolgedessen sinnvoll, diese Information getrennt von den eigentlichen Geoobjekten zu halten.

Die Erfahrungen mit GODOT haben gezeigt, daß der objektorientierte Ansatz für die Datenhaltung in geographischen Informationssystemen die Modellierung komplexer Anwendungen begünstigt, wie sie insbesondere im Umweltbereich die Regel sind, und eine vergleichsweise flexible Anpassung des Geoinformationssystems an anwendungsspezifische Anforderungen erlaubt.

6. Zeitliche Aspekte von Umweltdaten

Eine typische Anforderung an die Datenhaltung für Umweltinformationssysteme betrifft die Ablage von historischen Daten oder von Zeitreihen. Dies ist insbesondere wichtig, um Umweltveränderungen im zeitlichen Verlauf zu erkennen, zu dokumentieren und beurteilen zu

können, z.B. in Bezug auf Altlasten, Schadstoffimmission oder globale Umweltveränderungen (Global Change). Im Bereich der Flurbereinigung, Umweltplanung oder der Umweltverträglichkeitsprüfungen wird darüber hinaus auch die Systemeigenschaft gefordert, zukünftige, geplante Umweltzustände repräsentieren zu können, womöglich sogar in verschiedenen alternativen Versionen. Eine besondere Herausforderung, die am FAW im Rahmen des Forschungsprojekts CHRONOS (Temporale Aspekte in Geoinformationssystemen) zusammen mit der strässle Informationssysteme GmbH bearbeitet wird, besteht dabei darin, die zeitlichen Aspekte von Umweltdaten mit deren räumlichen Aspekten zu kombinieren und insbesondere die in diesem Umfeld erforderlichen zeitlich-räumlichen Interpretationen zu unterstützen [16]. Im Hinblick auf die Datenhaltung bedeutet dies, daß ein raum- und zeitbezogenes Datenmodell für geographische Daten entwickelt wird. In diesem Datenmodell hat jedes Geoobjekt einen Gültigkeitszeitraum, der durch die Zeitpunkte des Entstehens und des Vergehens des Umweltobjekts abgesteckt ist. Während dieses Gültigkeitszeitraums können sich die Attribute des Geoobjekts verändern, d.h. auch die Werte der Attribute haben einen Gültigkeitszeitraum und ein Attribut kann während der Lebensdauer eines Objekts sehr viele mögliche Werte aufweisen. Ein Beispiel hierfür ist die Veränderung der Bevölkerungszahl einer Stadt, das Vorkommen von Arten in einem Biotop oder die Wassergüteklasse eines Fließgewässers. Ein besonderes Problem ist die adäquate Repräsentation der Veränderung der Geometrie von Geoobjekten, etwa das Wachsen einer Stadt, das Schrumpfen eines Sees oder eines Waldgebietes oder die Änderung von Flußverläufen.

Obwohl CHRONOS konzeptionell ein objektorientiertes Datenmodell zugrundeliegt, hat es sich hier als vorteilhaft gezeigt, eine relationale Datenbank zur Implementierung des Datenmodells zu verwenden. Geoobjekte werden dabei durch eine Kennung, die sogenannte Objekt-ID, eindeutig identifiziert, die das Objekt über seine ganze Lebenszeit hinweg beibehält. Die verschiedenen Objektattribute inklusive der Geometrie sind auf unterschiedliche Relationen verteilt. Dies gilt auch für die Geometriedaten, die in sogenannten BLOBs (Binary Large Objects = „Binäre große Datenobjekte“) abgelegt sind. In diesen Relationen ist mit jedem Tupel die Objekt-ID des zugehörigen Objekts sowie ein Gültigkeitszeitraum vermerkt. Wenn nun der Zustand eines Objektes zu einem bestimmten Zeitpunkt ermittelt werden soll, so werden zunächst alle zu diesem Objekt gehörenden Tupel gesucht, deren Gültigkeitszeitraum den betrachteten Zeitpunkt umschließt. Ein Join über alle diese Tupel bezüglich der Objekt-ID des betrachteten Objekts ergibt dann den gesuchten Objektzustand zum fraglichen Zeitpunkt.

7. Abfragesprache

Die besondere Natur von Umweltdaten erfordert auch besondere Leistungsmerkmale für die Abfragesprache von Umweltdatenbanken. Objektorientierte Datenbanksysteme, die die Bildung von komplexen Objekten erlauben, erfordern reichhaltigere Sprachmittel als die herkömmlichen relationalen Systeme. Während Abfragen in relationalen Systemen stets aus der Gesamtheit aller Tupel in einer Relation schöpfen, können Abfragen in objektorientierten Datenbanksystemen gegen unterschiedliche Arten von Grundgesamtheiten gerichtet werden. So etwa gegen die Menge aller Instanzen einer Klasse (dies entspricht am ehesten dem relationalen Fall), aber auch gegen die Bestandteile eines komplexen Objekts oder gegen eine Menge von eigenständigen Objekten, die zu einem gegebenen Objekt in einer bestimmten Relation stehen. Da mit den beiden letzten Arten von Abfragen die Verbindungen in einem Netz von Objekten durchlaufen werden können, spricht man hier auch von navigatorischen Abfragen.

Eine weitere Besonderheit betrifft die Verwendung von räumlichen oder zeitlichen Prädikaten in Abfragen. So können aus einer Menge von Objekten all diejenigen Objekte selektiert

werden, die zu einem gegebenen Objekt in einer bestimmten räumlichen oder zeitlichen Beziehung stehen. Als räumliche Prädikate bieten sich hier alle binären geometrisch-topologischen Prädikate an, wie z.B. Überlappung, räumliches Enthaltensein oder eine vorgegebene Maximaldistanz zwischen zwei Objekten. Eine innerhalb eines gewissen vorgegebenen Rahmens vollständige Auflistung solcher Prädikate findet sich in [3]. Eine entsprechende Systematik für zeitliche Relationen findet sich in [2]. In Analogie zum Begriff „Spatialer Join“ [13] bezeichnen wir in unseren Arbeiten die allgemeinste Art der beschriebenen spatiotemporalen Abfragen als „Spatiotemporale Joins“.

Die Verwendung von spatiotemporalen Prädikaten in Anfragen bedeutet eine Erweiterung der herkömmlichen Anfragesyntax. Postrelationale, insbesondere objektorientierte Datenbanksysteme erlauben häufig die Verwendung von benutzerdefinierten Methoden oder Funktionen in einer Query; hier läßt sich die Erweiterung der Anfragesyntax durch Definition einer geeigneten Funktion vergleichsweise leicht durch den Benutzer vornehmen. Wenn dies nicht möglich ist, sind andere Techniken, z.B. die Verwendung von Präprozessoren für die Abfragebearbeitung erforderlich. Diese Technik, d.h. die Verwendung eines Präprozessors, kann im übrigen auch genutzt werden, um die Datenbankanfragesprache so zu erweitern, daß auch auf interpolierte Werte von Meßreihen in der gleichen Weise zugegriffen werden kann wie auf die explizit gespeicherten Meßwerte [17]. Hier ist allerdings dem Benutzer die der jeweiligen Interpolationsmethode zugrundeliegende Philosophie (Modellierung, Annahmen, Algorithmik) in geeigneter Form bekanntzumachen.

Ein besonderes Problem ist die Optimierung der Abarbeitung von spatiotemporalen Abfragen. Dies betrifft zum einen die Ablagestrukturen; denn in aller Regel kann eine schnellere Abfragebearbeitung erwartet werden, wenn zeitlich und räumlich benachbarte Objekte physisch möglichst nahe beieinander auf dem Speichermedium abgelegt werden. Zum andern betrifft dies die Verwendung räumlich-zeitlicher, also mehrdimensionaler Indexstrukturen, mit deren Hilfe alle Objekte aus einem bestimmten räumlichen Bereich und/oder einem bestimmten Zeitraum schnell aufgefunden werden können. Es gibt eine Vielzahl derartiger Indexstrukturen [11]. Das Problem ist jedoch, daß sich die wenigsten in handelsübliche Datenbanksysteme integrieren lassen. Vielversprechend ist hier das sogenannte Z-Ordering-Indizierungsschema, weil es sich auf die in Datenbanken üblichen eindimensionalen Indizes zurückführen läßt; dieses Schema findet auch im Rahmen des FAW-Projekts GODOT Anwendung [10].

8. Schemainformation

Eine besonders wichtige und generelle Art von Metainformation in Datenbanksystemen ist in der jeweiligen Schemainformation verkörpert. Solche Information findet sich beispielsweise in dem Data Dictionary eines relationalen Datenbanksystems. In einem objektorientierten Datenbanksystem besteht die Schemainformation aus den Beschreibungen aller Klassen der in der Datenbank gespeicherten Objekte. Solche Information läßt sich auf vielfältige Weise nutzen.

Wenn die Schemainformation in einer benutzergerechten Weise präsentiert wird, ist es dem Datenbankbenutzer möglich, sich über die in der Datenbank enthaltenen Klassen von Objekten zu informieren. Im Idealfall sind sogar Hilfe- und Dokumentationstexte Teil der Schemainformation und können dem Benutzer in geeigneter Form präsentiert werden. Wünschenswert ist es oft auch, daß der Benutzer diese Schemainformation – und damit die Struktur der Datenbank – verändern kann. Auf diese Weise wird es möglich, Schemamodifikationen vorzunehmen und die Datenbank an spezielle Bedürfnisse anzupassen.

Die Schemainformation kann von einem Umweltinformationssystem benutzt werden, um automatisch Präsentationen und Dialoge zu den in der Datenbank gespeicherten Objekte zu

generieren. Insbesondere ist das Vorhandensein entsprechender Metainformation entscheidend für die Möglichkeit der Verwendung natürlicher Sprachen für die Benutzerinteraktion [5]. Wenn die Schemainformation entsprechend reichhaltig angelegt ist, kann sie auch genutzt werden, um Umsetzungsprogramme zu steuern, die die in der Datenbank gespeicherten Daten in Externformate umsetzen oder umgekehrt. Schemainformation ist also ein wertvoller Input, um Heterogenität zu überwinden; dies erleichtert beispielsweise die Integration heterogener Datenbanken [8].

9. Ausblick

Das Thema der zentralen Bedeutung der Verfügbarmachung aller zur korrekten Interpretation von Umweltdaten erforderlichen Zusatzinformation ist bislang zu wenig in das Zentrum der Aufmerksamkeit der Entscheidungsträger und der tangierten wissenschaftlichen Disziplinen gerückt [14]. Schlimmer noch: es wird in den nationalen und internationalen Förderprogrammen zur Umweltthematik bis heute nicht wirklich deutlich, daß die Bewältigung der Daten- und Modellproblematik oft der Schlüssel und die unbedingte Voraussetzung zum Erfolg ist und die jeweiligen Fachwissenschaften alleine hier nicht den Erfolg garantieren können. Aus diesem Grund ist es nicht verwunderlich, daß es bis heute noch wenige Beispiele für Umweltinformationssysteme gibt, die auf Umweltdatenbanken zurückgreifen können, in denen neben den „eigentlichen“ Daten auch alle erforderliche Zusatzinformationen, also Metainformationen im weitesten Sinn, abgelegt sind. In der jüngsten Zeit sind jedoch erste Fortschritte bei der informationstechnischen Behandlung von Metainformation zu beobachten.

So ist die Repräsentation von Metainformation zu unterschiedlichen Umweltdatenbanken und Umweltinformationssystemen Gegenstand des Projekts HEM (Harmonization of Environmental Measurement [15]). Dies ist eine Reaktion darauf, daß ohne derartige Metainformation der Schatz der international vorhandenen Daten nicht genutzt werden kann. Der Umweltdatenkatalog (UDK) [22], dessen Einführung und Weiterentwicklung derzeit in einer Kooperation deutscher Bundesländer unter der Führung des Landes Niedersachsen vorangetrieben wird und für den sich auch die Republik Österreich entschieden hat, zählt zu den ersten sogenannten Metainformationssystemen, die ausschließlich Metadaten enthalten. Mit dem UDK ist es möglich, sich über die verschiedenen Datensammlungen zu informieren, die in der Umweltverwaltung eines Bundeslandes vorgehalten werden. Was der UDK als isoliert ablaufendes System jedoch noch nicht leisten kann, ist die Unterstützung beim Zugriff auf die gefundenen Datenbestände. Hier sind dann noch enorme technische Probleme in der Überwindung von Heterogenität (Datenbanksysteme, Betriebssysteme, Sprachen, Netzwerke usw.) zu lösen.

Metainformation, so zeigt sich, ist wahrscheinlich das einzige Mittel, das uns bleibt, um mit der Heterogenität und Verteiltheit von Umweltdaten umzugehen [7]. In unserer Welt, in der elektronische Kommunikationsnetze, „informationstechnische Nervensysteme“, zunehmend an Bedeutung gewinnen, ist die Verfügbarkeit von Metainformation eine notwendige, wenn auch noch nicht hinreichende Voraussetzung, um die notwendige Interoperabilität von Informations- und Datenbanksystemen sicherzustellen und damit ein grenzenüberschreitendes, kooperatives Handeln bei der Bewahrung der natürlichen Ressourcen der Erde zu ermöglichen. Daß bei einer eigentlich gebotenen weltweiten Perspektive hier noch gewaltige Herausforderungen anstehen, ist klar, wenn man alleine die technischen Schwerpunkte berücksichtigt, mit denen sich beispielsweise ein so fortschrittliches Konzept wie das Umweltinformationssystem (UIS) Baden-Württemberg – erfolgreich – auseinandergesetzt hat. Zu hoffen ist, daß etwa die geplante Europäische Umweltagentur für unseren Kontinent wesentliche Fortschritte in diesem Umfeld erreichen kann.

Literatur

- [1] Arbeitsgemeinschaft der Vermessungsverwaltungen der Länder der Bundesrepublik Deutschland (AdV): Amtliches Topographisch-Kartographisches Informationssystem ATKIS, Bezug über das Landesvermessungsamt NRW. Bonn 1989.
- [2] Allen, J.F.; Hayes, P.J.: A Common-Sense Theory of Time. In: *Proc. IJCAI*. Los Angeles, Ca., August 1985.
- [3] Bähr, U.: *Untersuchungen zu räumlichen Abfragesprachen*. Diplomarbeit. Technische Universität München, 1992.
- [4] Baumewerd-Ahlmann, A.; Zink, L.: Umweltdatenbanken – Sachstand und Zukunftsperspektiven. In [18], S. 79 - 101.
- [5] Becker, R.; Küpper, D; Strobel, M.; Rösner, D: A Cooperative Natural Language Information System. In: *Information Processing 92; Proc. of the 12th World Computer Congress*. Vol. II. Elsevier, 1992.
- [6] Special Section: Next Generation Database Systems. *Communications of the ACM*. 34(10):30-120, October 1991.
- [7] Denzer, R.; Güttler, R.: Integration von Umweltdaten. In [18], S. 227 - 240.
- [8] Endrikat, A.; Michalski, R.: The WINHEDA Prototype: Knowledge-based Access to Distributed Heterogeneous Knowledge Sources. In: Schader, M. (Hrsg.): *Analyzing and Modeling Data and Knowledge*. Berlin – Heidelberg – New York: Springer-Verlag, 1991.
- [9] Ebbinghaus, J.; Günther, O.; Hess, G.; Lamberts, J.; Riekert, W.-F.: Objektorientierte Datenbanksysteme für Geo- und Umwelthanwendungen – eine Vergleichsstudie. In: *3. Internationales Anwenderforum Geoinformationssysteme*, Tagungsband, Duisburg, März 1993.
- [10] Gaede, V.; Riekert, W.-F.: Spatial Access Methods and Query Processing in the Object-Oriented GIS GODOT. FAW-Forschungsbericht, Ulm, 1994.
- [11] Günther, O.: *Efficient Structures for Geometric Data Management*. Lecture Notes in Computer Science No. 337, Berlin – Heidelberg – New York: Springer-Verlag, 1988.
- [12] Günther, O.; Kuhn, H.; Mayer-Föll, R.; Radermacher, F.J. (Hrsg.): *Konzeption und Einsatz von Umweltinformationssystemen*. Berlin – Heidelberg – New York: Springer-Verlag, 1991.
- [13] Günther, O.: Efficient Computation of Spatial Joins. FAW-Forschungsbericht FAW-TR-93008. Ulm, Juli 1993.
- [14] Jaeschke, A.; Keitel, A.; Mayer-Föll, R.; Radermacher, F.J.; Seggelke, J.: Metawissen als Teil von Umweltinformationssystemen. In: [12], S. 115-130.
- [15] Keune, H.; Murray, A.B.; Benking, H.: Harmonization of Environmental Measurement. *Geo-Journal* 23/3:249-255. Dordrecht, Boston: Kluwer Academic Publishers 1991.
- [16] Luttermann, H.: *Marktstudie und Sachstandsbericht zum Thema „Zeit in geographischen Informationssystemen“*. FAW-B-94001, FAW Ulm, Februar 1994.
- [17] Neugebauer, L.: Einbettung von Interpolationsverfahren in die Anfragesprache SQL zur Bearbeitung von Umwelt-Meßwerten. In: [12], S. 269-288.
- [18] Page, B.; Hilty, L.M. (Hrsg.): *Umweltinformatik. Informatikmethoden für Umweltschutz und Umweltforschung*. München – Wien: R. Oldenbourg Verlag 1994.
- [19] Radermacher, F.J.: The Importance of Metaknowledge for Environmental Information Systems. In: Günther, O.; Schek, H.-J. (eds.): *Large Spatial Databases. Proceedings*. S. 35-44. Berlin – Heidelberg – New York: Springer 1991.
- [20] Riekert, W.-F.: Das Forschungsprojekt GODOT: Geodatenhaltung mit objektorientierten Techniken. In: *3. Internationales Anwenderforum Geoinformationssysteme*, Tagungsband, Duisburg, März 1993.
- [21] Riekert, W.-F.; Henning, I.; Schmidt, F.: Integration von heterogenen Komponenten des Umweltinformationssystems (UIS) Baden-Württemberg. In: *2. Workshop „Integration von Umweltdaten“*, KfK-5314, Kernforschungszentrum Karlsruhe, 1994.
- [22] Schütz, T.; Lessing, H.: Metainformationen von Umwelt-Datenobjekten – Zum Datenmodell des Umwelt-Datenkataloges Niedersachsen. In: Jaeschke, A.; Kämpke, T.; Page, B.; Radermacher, F. (Hrsg.): *Informatik für den Umweltschutz; Proc. 7. Symposium*. S. 19-28. Berlin – Heidelberg – New York: Springer 1993.

Prof. Dr. Dr. Franz Josef Radermacher ist seit 1987 Leiter des Forschungsinstituts für anwendungsorientierte Wissensverarbeitung (FAW), Ulm und Professor für Informatik an der Universität Ulm. Von 1988-1992 war er Präsident der Gesellschaft für Mathematik, Ökonomie und Operations Research (GMÖOR); 1990 Gründungssprecher der Arbeitsgemeinschaft der deutschen KI-Institute (AKI).

Dr. Wolf-Fritz Riekert ist seit 1993 Leiter des Anwendungsbereichs Umweltinformationssysteme im Forschungsinstitut für anwendungsorientierte Wissensverarbeitung (FAW), Ulm. Seit 1989 leitete er verschiedene FAW-Projekte, u.a. auf den Gebieten der Auswertung von Satellitenbilddaten, der Geoinformationssysteme, der objektorientierten Datenbanken und der Systemintegration.