

Übersichtsbeiträge

Umweltrelevante Datenbanken

– Online-Datenbanken und CD-ROMs

¹Kristina Voigt, ¹Rainer Brüggemann, ²Joachim Benz

¹ GSF – Forschungszentrum für Umwelt und Gesundheit GmbH, Projektgruppe Umweltgefährdungspotentiale von Chemikalien (PUC), Ingolstädter Landstraße 1, D-85758 Oberschleißheim

² Gesamthochschule Kassel, Fachbereich 20, Nordbahnhofstraße 1A, D-37213 Witzenhausen

Zusammenfassung

Datenbanken zu Thema Umweltschutz finden zunehmendes Interesse in der Fachwelt und der Öffentlichkeit. Große Datenbankanbieter (Hosts) haben dies erkannt und unterstützen die Suche nach umweltrelevanten Informationen. Im Informationssystem Umweltchemikalien werden seit einigen Jahren in der GSF – Forschungszentrum für Umwelt und Gesundheit GmbH im Auftrag des Bayerischen Staatsministeriums für Landesentwicklung und Umweltfragen Metadatenbanken im Fachgebiet „Umweltschutz und Chemikalien“ entwickelt und gepflegt. Die Bewertung der Datenbanken stellt einen wichtigen neuen Forschungsschwerpunkt dar. Dabei werden Hasse-Diagramme zur Anwendung gebracht.

Schlagwörter: Umweltrelevante Datenbanken; Online-Datenbanken; CD-ROMs; Hosts; Cluster; Umwelt-Cluster; Hasse-Diagramme; Bewertungsverfahren; Scoring Systeme

1 Einleitung

Mit der rasch ansteigenden Anzahl der Publikationen in Wissen und Technik ist ebenfalls eine Zunahme der Datenbanken auf diesem Gebiet verbunden. In der weltweit größten Zusammenstellung von Datenbanken (Online-Datenbanken und CD-ROMs), dem Gale Directory of Databases 1994 [1], sind 8 400 Datenbanken beschrieben. Es handelt sich dabei um 5 200 Online-Datenbanken und 3 200 Datenbanken auf CD-ROM, Disketten etc. Das Gale Directory of Databases hat die Standardwerke Computer-Readable Databases, Cuadra/Gale's Directory of Online Databases und Directory of Portable Databases vereinigt. Das Gale Directory wird sowohl als zweibändiges Buch als auch als Online-Datenbank bei DIALOG sowie auf CD-ROM bei Silver Platter angeboten. Die Auswertung der vorherigen Auflage ergab einen Prozentsatz von 20 für die Themenschwerpunkte Naturwissenschaft, Technik, Medizin [2].

2 Datenbanken für Chemikalien und Umweltschutz

Es gibt international und national einige Absätze aus der Gesamtheit der Datenbanken diejenigen herauszuarbeiten, die sich mit der Thematik des Umweltschutzes auseinandersetzen. Exemplarisch sind hier zu nennen: die Zusammen-

stellung von Umweltdatenbanken der britischen Association for Information Management [3], der Umweltdatenbankführer [4], die Datenquellen für Umweltchemikalien [5].

Auch werden in zunehmendem Maße Seminare und Kurse über umweltrelevante Datenbanken abgehalten. Beispielsweise sei an dieser Stelle auf Symposien über umweltbezogene Gefahrstoffdatenbanken [6] und Vorlesungen im Weiterbildungsstudiengang „Umweltrecht und Umweltökonomie“ [7] hingewiesen.

Die Datenbankanbieter (Hosts) sind zunehmend an der Thematik des Umweltschutzes interessiert. In den hostspezifischen Datenbankgruppen, die man Superbase oder auch Cluster nennt, existiert die Gruppe „Environment“ bei vielen Datenbank Anbietern. So nennt z.B. der Host DIMDI seine Gruppe Umweltschutz, Umweltmedizin [8], DIALOG [9] und STN [10] ihre jeweiligen Cluster Environment.

In der **Tabelle 1** sind 54 als umweltrelevant eingestufte Datenbanken von den drei ausgewählten Hosts DIALOG, DIMDI und STN aufgelistet. In der letzten Spalte ist unsere noch später näher beschriebene Metadatenbank DADB aufgeführt. Das Cluster von STN umfaßt 28, das von DIALOG 24 und das von DIMDI 18 Online-Datenbanken. Es wird anhand der **Tabelle 1** deutlich, daß nur einige wenige Datenbanken bei allen drei Hosts gleichzeitig aufliegen. Es handelt sich hierbei um BIOSIS Previews, CAB Abstracts und EMBASE. Es fällt auf, daß beispielsweise die Datenbank MEDLINE, die von allen drei betrachteten Host verfügbar gemacht wird, bei DIALOG nicht im Environment Cluster aufgeführt wird. Diese Datenbank wird von DIMDI und STN als umweltrelevante Datenbank angesehen, nicht aber von DIALOG. Ebenso wird die Datenbank HSDB von STN aufgelistet nicht aber von DIMDI, obwohl diese Datenbank auch im Angebot dieses Hosts zu finden ist. In diesem Zusammenhang ist anzumerken, daß sich überwiegend bibliographische Umweltdatenbanken in den ausgewerteten Umweltclustern befinden und kaum Volltext- und Faktendatenbanken.

Die für Umweltfragen wichtige Datenbank ECDIN, die bei DIMDI angeboten wird, taucht jedoch im Umweltcluster dieses Hosts nicht auf. Die Realisierung der Vereinigung von mehreren Datenbanktypen (z.B. Fakten- und biblio-

Tabelle 1: Cluster „Environment“ (Umwelt) bei drei ausgewählten Hosts

Nr.	Datenbank	Host	Host	Host	Metadb.
1	ABI-INFORM			STN	DADB
2	AGRICOLA		DIMDI		DADB
3	AGRIS		DIMDI		DADB
4	APILIT			STN	
5	AQUASCI			STN	DADB
6	ASFA	DIALOG	DIMDI		DADB
7	BALTIC		DIMDI		DADB
8	BIOBUSINESS			STN	DADB
9	BIOLIS		DIMDI		
10	BIOSIS	DIALOG	DIMDI	STN	DADB
11	BNA Daily News	DIALOG			
12	CA	DIALOG		STN	DADB
13	CAB	DIALOG	DIMDI	STN	DADB
14	CAPreviews			STN	
15	CBNB			STN	DADB
16	CEN			STN	DADB
17	CIN			STN	DADB
18	CJACS			STN	DADB
19	COMPENDEX	DIALOG		STN	DADB
20	CRIS/USDA	DIALOG			DADB
21	CSNB			STN	DADB
22	DIALOG Source Eng.	DIALOG			
23	ELFIS		DIMDI		DADB
24	EMBASE	DIALOG	DIMDI	STN	DADB
25	ENERGIE			STN	
26	ENERGY			STN	DADB
27	Energy S.T.	DIALOG			DADB
28	ENVIROLINE	DIALOG	DIMDI		DADB
29	Environm. Bib.	DIALOG			DADB
30	Fed. News Ser.	DIALOG			
31	Fed. Register	DIALOG			DADB
32	Geoarchive	DIALOG			DADB
33	Geobase	DIALOG			DADB
34	GEOREF			STN	DADB
35	HEALSAFE			STN	DADB
36	HSDB			STN	DADB
37	IPA		DIMDI		DADB
38	JICST-E			STN	DADB
39	LIFESCI			STN	DADB
40	MEDLINE		DIMDI	STN	DADB
41	NTIS	DIALOG		STN	DADB
42	OCEAN	DIALOG		STN	DADB
43	Paperchem	DIALOG			DADB
44	PIRA	DIALOG			DADB
45	POLLUTION	DIALOG		STN	DADB
46	PSTA		DIMDI		DADB
47	PTS Newsl. Db.	DIALOG			
48	Public Opinion	DIALOG			
49	SCISEARCH		DIMDI		DADB
50	SOMED		DIMDI		DADB
51	TOXLINE	DIALOG	DIMDI		DADB
52	TOXBIO		DIMDI		DADB
52	TOXCAS		DIMDI		DADB
53	UFORDAT			STN	DADB
54	ULIDAT			STN	DADB

graphische Datenbanken) in einem Cluster scheint wegen des unterschiedlichen Aufbaus der Datenbanken noch Schwierigkeiten zu bereiten. Es wäre jedoch möglich und technisch zu realisieren, einen Faktendatenbanken- und einen bibliographischen Datenbanken-Cluster zu erstellen. Aus Tabelle 1 wird ferner deutlich, daß die Hosts nicht nur unterschiedliche umweltrelevante Datenbanken in ihrem Angebot haben, sondern auch verschiedene Einstufungen

in diesem Zusammenhang vornehmen. Der Nutzer des Umweltclusters kommt bei jedem Host also zu deutlich unterschiedlichen Rechercheergebnissen, wenn er nach Umweltschutzaspekten sucht.

Seit einigen Jahren arbeiten wir im Auftrag des Bayerischen Staatsministeriums für Landesentwicklung und Umweltfragen an dem Forschungsschwerpunkt Daten und Informationen über Chemikalien und Umweltschutz. In diesem Zu-

sammenhang haben wir Datenquellen in bezug auf die Thematik des Umweltschutzes im Zusammenhang mit Chemikalien ausgewertet. Dazu haben wir sog. Metadatenbanken, d.h. Datenbanken der Datenquellen, aufgebaut. Folgende Metadatenbanken seien genannt:

DALI: Metadatenbank der Literatur

DADB: Metadatenbank der Online-Datenbanken

DACD: Metadatenbank der CD-ROMs [11, 12].

Der Dokumentenstand (Februar 1994) beträgt für DALI: 915, für DADB: 409 und für DACD: 247 Dokumente, d.h. 656 Datenbanken (Online-Datenbanken und CD-ROMs) enthalten Informationen über das Thema „Chemikalien und Umweltschutz“.

Aus dem bisher Dargestellten ergibt sich die Anzahl von 656 Datenbanken mit Umweltbezug (→ Tabelle 2).

Tabelle 2: Auswahl von Datenbanken mit Umweltbezug

Gesamtheit der Datenbanken	8400
Naturwissenschaft, Technik	2100
Datenbanken mit Umweltbezug	656

In unserer Metadatenbank der Online-Datenbanken haben wir von den in Tabelle 1 aufgeführten 54 umweltrelevanten Online-Datenbanken 45 inkorporiert. Diejenigen Datenbanken, die sich mit Nachrichten befassen oder nicht uneingeschränkt verfügbar sind, werden im allgemeinen nicht von uns ausgewertet.

3 Metadatenbank DADB und DACD

Neben den oben aufgeführten Metadatenbanken haben wir im Informationssystem Umweltchemikalien eine Kommunikationssoftware mit einem Suchhilfesystem entwickelt [13]. In Abb. 1 sind die Bausteine des Forschungsprojektes aufgeführt.

3.1 Hosts und deren Online-Datenbanken in DADB

Datenbanken der folgenden Host sind inkorporiert: BRS, CIS, DATASTAR, DIALOG, DIMDI, ECHO, ESA, FIZ Technik, GBI, ICSTI, ORBIT, Questel, STN, TDS Numerica, VINITI und einige kleinere Hosts, die nur ein oder zwei Datenbanken anbieten. Mit Ausnahme der beiden großen russischen Hosts ICSTI und VINITI haben wir Zugriff zu allen Datenbankanbietern.

3.2 Typisierung von Online-Datenbanken in DADB

Neben der üblichen Einteilung der Online-Datenbanken in bibliographische (BI), Fakten- oder numerische (NU) und Volltextdatenbanken (VT) unterscheiden wir gemäß unserer Thematik noch Reaktionsdatenbanken (RD), Chemikalien-Datenbank-Verzeichnisse (CN), Hersteller-Verzeichnisse (DI), Forschungsdatenbanken (FD) und Metadatenbanken (MD). In Reaktionsdatenbanken werden Informationen über Reaktionen von chemischen Substanzen be-

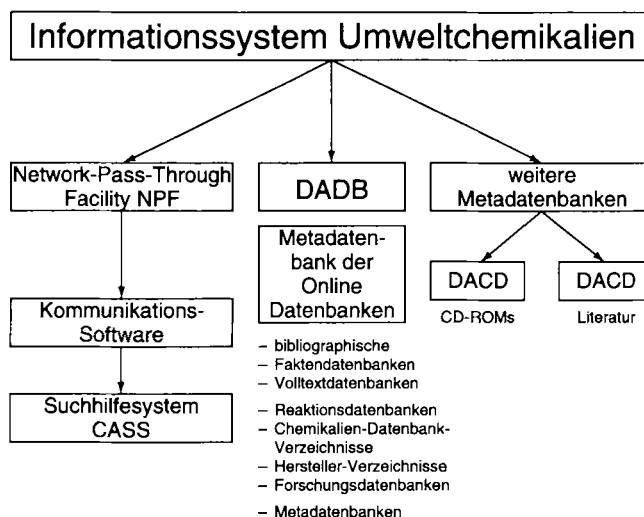


Abb.1: Überblick über das Informationssystem Umweltchemikalien

trachtet. Chemikalien-Datenbank-Verzeichnisse sind Datenbanken, die Identifikationsmerkmale zu Chemikalien liefern und ggfs. auch auf andere Datenbanken verweisen, in denen sich weitere Informationen über die gewünschten Substanzen befinden. Beispiele hierfür sind CHEMSEARCH bei DIALOG und CHEMNAME bei DIMDI. Hersteller-Verzeichnisse umfassen in der Regel nicht nur den Namen und die Adresse der Hersteller, sondern haben meist auch mehr oder weniger umfangreiche Datenprofile über die hergestellten Chemikalien. In Forschungsdatenbanken findet man Beschreibungen von Forschungsvorhaben mit Angaben von Forschungsnehmern und Forschungsgebern. Eine weitere Art von Online-Datenbanken bilden die Datenbanken über Datenquellen, die sogenannten Metadatenbanken. Bekannte Beispiele sind CUADRA Directory of Online Databases (Hosts: DataStar, DIALOG) und Gale Directory of Databases (Host: DIALOG).

Unter unseren 409 ausgewählten Datenbanken sind:

- 211 bibliographische Datenbanken
- 100 Faktendatenbanken (numerische Datenbanken, Abschätzprogramme)
- 32 Volltextdatenbanken
- 4 Reaktionsdatenbanken
- 12 Chemikalien-Datenbank-Verzeichnisse
- 15 Hersteller-Verzeichnisse
- 18 Forschungsdatenbanken
- 17 Metadatenbanken

Für unsere Auswertung ziehen wir zunächst nur die bibliographischen, Fakten- und Volltextdatenbanken sowie Chemikalien-Datenbank-Verzeichnisse heran.

In Tabelle 3 wird die Aufteilung der Datenbank-Typen auf die verschiedenen Hosts gezeigt.

STN und DIALOG haben die größte Anzahl von Umweltchemikalien-relevanten Datenbanken in ihrem Angebot, gefolgt von DataStar. CIS, DIMDI und ESA verfügen ebenso über eine große Anzahl von relevanten Datenbanken. Die Hosts CIS und TDS Numerica verfügen über eine große Anzahl von Faktendatenbanken auf dem zu betrach-

Tabelle 3: Aufteilung der Umweltchemikalien-relevanten Datenbanken auf die verschiedenen Hosts

Host	Anzahl DB	BI	NU, VT, CN
BRS	20	10	8
CIS	34	4	29
DataStar	43	29	12
DIALOG	54	32	15
DIMDI	46	30	13
ECHO	9	1	–
ESA	23	22	1
GBI	8	1	1
ICSTI	11	11	–
ORBIT	26	18	5
Questel	13	9	3
STN	89	41	32
TDS	15	–	12
VINITI	6	3	1
andere	12	–	12
	409	211	144

teten Spezialgebiet. Da Wirtschaftsinformationen über Chemikalien und die sogenannten Metadatenbanken (Datenbanken über Datenbanken) für uns von zunehmendem Interesse sind, haben wir auch den Host GBI, als Anbieter von vornehmlich wirtschaftsorientierten Datenbanken, angemietet. In dieser Tabelle zeigt sich auch deutlich, daß die meisten der betrachteten Datenbanken, d.h. ca. 60% bibliographische sind, dagegen nur ca. 35% Fakten- oder Volltext-Datenbanken und Chemikalien-Datenbank-Verzeichnisse. Das Überwiegen von bibliographischen gegenüber Fakten-, Volltext-Datenbanken und Chemikalien-Datenbank-Verzeichnissen findet man mit Ausnahme von CIS und STN bei allen betrachteten Host.

3.3 Datenbankdefinition von DADB und DACD

Die vollständige Datenbankdefinition von DADB und DACD wurde bereits in verschiedenen Publikationen festgehalten und beschrieben [11, 12, 14]. An dieser Stelle sei nur auf diejenigen Datenfelder verwiesen, welche die Auswertung der Datenbanken in Bezug auf die Thematik „Umweltschutz und Chemikalien“ hervorheben. Diese sogenannten umweltrelevanten Datenfelder sind in Tabelle 4 aufgeführt. Sie sind in beiden Metadatenbanken DADB und DACD dieselben.

Tabelle 4: Umweltrelevante Datenfelder in DADB und CACD

Pos.	Feldname	Kurzb.	Feldwerte
26	Informationstyp	IT	Alphanum.
27	Anwendung der Chemikalien	AW	Alphanum.
28	Anzahl der Chemikalien	AN	ganze Zahl
29	Deskriptoren	DE	Alphanum.
30	Testset	TE	ganze Zahl

Unter dem Begriff „Informationstyp“ oder „Informationsparameter“ verstehen wir beispielsweise physikalisch-chemische Eigenschaften oder Strukturmerkmale. Es gibt Online-Datenbanken, deren Schwerpunkte auf einzelnen Informationstypen liegen. Ebenso sind einige Datenbanken

für spezielle Anwendungen von Chemikalien (z.B. Pestizide) ausgelegt. Darüber hinaus dient das Datenfeld „Anzahl der Chemikalien“ als grobe Einstufung der Bedeutung der entsprechenden Online-Datenbanken für unser Thema. Es gibt Datenbanken, die über 100 000 Chemikalien-Informationen vorhalten; bei anderen Quellen liegt die Anzahl der Stoffe bei einigen Hundert. Für das wichtigste inhaltserschließende Datenfeld, das „Deskriptor“-Feld, wurde ein Thesaurus entwickelt, der ca. 200 Schlagwörter oder Suchbegriffe aus den Fachgebieten des Umwelts-, Gesundheits- und Arbeitsschutzes enthält. Besondere Schwerpunkte nehmen die umweltrelevanten Parameter, wie das Vorkommen in den verschiedenen Umweltmedien und Ökotoxizitätsparameter ein. Im Zusammenhang mit den weiteren Forschungsschwerpunkten in der Projektgruppe Umweltgefährdungspotentiale von Chemikalien (PUC) sind Datenquellen, die Informationen über physikalisch-chemische Eigenschaften enthalten, von besonderem Interesse [15]. Das „Testset“-Feld dient zur Darstellung, welcher Stoff aus unserer Testchemikalien-Liste mit 68 umweltrelevanten Chemikalien in der entsprechenden Online-Datenbank enthalten ist [16].

4 Ansätze zur Bewertung von Datenbanken

Der Nutzer der Metadatenbanken DADB und DACD erfährt also, wo welche Informationen über die Thematik Umweltschutz und Chemikalien zu finden sind. Es ist also ein pragmatisches Hilfsmittel für die Auswahl der geeigneten Datenbank(en). Zusätzlich zu diesen Informationen aus DADB und DACD werden wegen der rasch anwachsenden Zahl der Datenquellen auch Konzepte zur Bewertung von Chemikalien-relevanten bzw. umweltrelevanten Datenbanken dringend benötigt. In einem weiteren tiefergreifenden Schritt im Forschungsprojekt Informationssystem Umweltchemikalien entwickeln wir einen Ansatz eines Scoring Systems für Datenbanken. In vielen Fällen ist es notwendig und sinnvoll, die für gewisse Randbedingungen „geeigneten“ Datenbanken für eine spezielle Fragestellung auszuwählen. Die Notwendigkeit der Einstufung bzw. Bewertung von Datenbanken wird auch im Fachgebiet der Informationswissenschaft erkannt. Man beginnt hier mit der Evaluierung der Datenbanken nach allgemeinen Kriterien der Informationswissenschaft. Hier sei auf einen Ansatz für die Bewertung von CD-ROMs nach den Kriterien Konsistenz, effiziente Ausnutzung der Recherchezeit, Einfachheit der CD-ROM, Objektivität der Beschreibung der CD-ROM, Flexibilität der Abfrage verwiesen [17, 18].

Im Gegensatz zu diesen mehr allgemeinen Bewertungsgrundsätzen entwickeln wir ein Konzept, das neben allgemeinen Datenbank-Bewertungskriterien auch solche enthalten, die auf potentielle Anwender-Fragestellungen zugeschnitten sind. Einerseits enthält es ein sog. Scoring-System für Datenbanken, andererseits sollen die Bewertungen möglichst die Vielzahl der Kriterien berücksichtigen, die das zu bewertende Objekt (Datenbank) aufweist. Somit ist nach mathematischen Verfahren Ausschau zu halten, die die „Parallelität von Informationen“ möglichst beinhalten. Eine Möglichkeit bildet die mathematische Disziplin der

Verbandstheorie und die Visualisierung von Ordnungsrelationen durch die Technik der Hasse-Diagramme. Hasse-Diagramme werden zum Teil in unserer Projektgruppe Umweltgefährdungspotentiale von Chemikalien (PUC) in der GSF-Forschungszentrum für Umwelt und Gesundheit GmbH angewendet; ein Programm für Personal-Computer, Betriebssystem DOS 3.0 und aufwärts wird derzeit entwickelt [19].

Die Informationsgewinnung aus der Vielzahl unterschiedlich strukturierter Datenquellen kann nach verschiedenen Gesichtspunkten, d.h. nach verschiedenen Bewertungskriterien, optimiert werden. Es werden daher z.Zt. erste Ansätze für ein Ranking bzw. für eine Bewertung der Datenquellen erarbeitet.

Bewertungskriterien für Datenquellen lassen sich grob in folgende drei Gruppen einteilen:

- Bewertungen, basierend auf der Charakterisierung von Umweltchemikalien
- Bewertungen, basierend auf umweltrelevanten Datentypen
- Bewertungen der Datenquellen allgemein.

Für die Auswahl der Kriterien spielen neben fachlichen Gesichtspunkten, selbstverständlich auch allgemeine Datenbank-Parameter eine Rolle: Kosten der Datenbanken, die Frequenz der Updates, der Umfang der Online-Datenbanken, die Verfügbarkeit der Online-Datenbanken (bei einem oder bei mehreren Hosts), das Vorhandensein der Daten auf/in anderen Medien (CD-ROMs, gedruckter Form).

4.1 Nutzung der Hasse-Diagramme zur Bewertung von Datenbanken

Die Hasse-Diagramme visualisieren sog. „Halbordnungen“ (nicht strikte Ordnungen), die z.B. entstehen, wenn Objekte nicht nur durch ein Merkmal gekennzeichnet werden. Die zugrundeliegende Theorie wird als ein Teil der Verbandstheorie betrachtet. Zu den wichtigsten Anwendungen von Hasse-Diagrammen im hier dargelegten Kontext gehört das Ranking von Objekten und die Feststellung von verträglichen Anordnungsfolgen.

Die Technik der Hasse-Diagramme wird bereits in verschiedenen Bereichen des Umweltschutzes angewendet:

- Einstufung von Chemikalien nach ihrer Umweltgefährlichkeit [20, 21, 22]
- Vergleichende Bewertung von Abfalldeponien [23]
- Vergleich ökologisch-mathematischer Modelle [24]
- Validierung von Quantitativen Strukturwirkungsbeziehungen [25]
- Vergleichende Bewertung von Flußabschnitten [26]
- Vergleichende Bewertung von aquatischen Wirkungstests [27]

Ebenso haben wir begonnen, Datenbanken unter Einsatz von Hasse-Diagrammen zu bewerten. Folgende Absätze sind erarbeitet worden:

- Einstufung von Chemikaliendatenbanken nach allgemeinen Bewertungskriterien [28]

- Einstufung von Umweltdatenbanken nach allgemeinen Bewertungskriterien [29]
- Vergleichende Bewertung von Chemikaliendatenbanken nach umweltrelevanten Bewertungskriterien [30].

Da die Auswahl der Bewertungskriterien einen Einfluß auf die Struktur von Hasse-Diagrammen hat, können wichtige oder weniger bedeutende Kriterien für die Einstufung von Datenbanken selektiert werden. In diesem Sinne unterstützt die Technik der Hasse-Diagramme auch das Ziel möglichst scharfe aussagekräftige Kriterien zu entwickeln.

4.2 Hasse-Diagramm für ausgewählte Umweltdatenbanken

Aus der **Tabelle 1** werden 14 Online-Datenbanken (Objekte), die hervorgehoben sind anhand von Bewertungsziffern (Scores) charakterisiert und das daraus resultierende Hasse-Diagramm vorgestellt und diskutiert. Es handelt sich um diejenigen Datenbanken, die in zwei oder drei Clustern auftreten und den beiden deutschen Umweltdatenbanken ULIT (ULIDAT) und UFOR (UFORDAT), die zwar von STN, FIZ Technik und DataStar, nicht aber von DIMDI und DIALOG angeboten werden.

Als „Evaluationskriterien“ werden folgende fünf Kriterien herangezogen:

1. Anzahl der Chemikalien
2. Identifikationsmerkmale
3. Informationsparameter für Umweltchemikalien
4. Art der Chemikalien
5. Validität/Güte der Online-Datenbanken

Das Bewertungsverfahren (Scoring) wird nach einem sechs Zahlen umfassenden Scoring System durchgeführt (→ **Tabelle 5**).

Tabelle 5: Bewertungsziffern (Scores) und deren Bedeutung

5 = sehr gut	2 = ausreichend
4 = gut	1 = mangelhaft
3 = befriedigend	0 = ungenügend

Auf die Vergabe der Scores für die einzelnen Kriterien wird hier nicht näher eingegangen, da die Vorgehensweise bereits an anderer Stelle diskutiert und publiziert wurde [30] (→ **Tabelle 6**).

Das korrespondierende Hasse-Diagramm für 14 Objekte (Online-Datenbanken) in bezug auf die Bewertung durch die fünf Evaluationskriterien (Attribute) Anzahl der Chemikalien, Identifikationsmerkmale, Informationsparameter für Umweltchemikalien, Art der Chemikalien, Validität/Güte der Online-Datenbanken wird in **Abb. 2** dargestellt.

Das Diagramm zeigt, daß es keine „beste“ Online-Datenbank (kein größtes Objekt in der Terminologie der Hasse-Algebra) gibt. Das gezeigte Ranking von 14 exemplarisch ausgewählten Online-Datenbanken, die in bezug auf Umweltschutz relevant sind, stellt die Online-Datenbanken TOXLINE (TO) und Chemical Abstracts (CA) als besonders hoch eingestufte Datenbanken heraus. Man spricht in der Hasse-Terminologie in diesem Zusammenhang von

Tabelle 6: Online-Datenbanken (Objekte) und deren Scores

Abk.	Objekt	AN	ID	IP	AR	QU
AS	ASFA	1	1	4	1	3
BI	BIOSIS	3	2	5	1	3
CA	Chem. Abs.	5	2	5	1	3
CB	CAB	2	1	3	1	3
CP	COMPENDEX	3	1	4	1	3
EM	EMBASE	3	2	5	1	2
EN	ENVIROLINE	1	1	2	1	3
ME	MEDLINE	2	2	4	1	2
NT	NTIS	1	2	4	1	2
OC	OCEAN	1	1	3	1	3
PO	POLLUTION	2	1	5	4	3
TO	TOXLINE	4	3	5	4	3
UF	UFORDAT	0	1	5	1	1
UL	ULIDAT	2	1	5	1	1

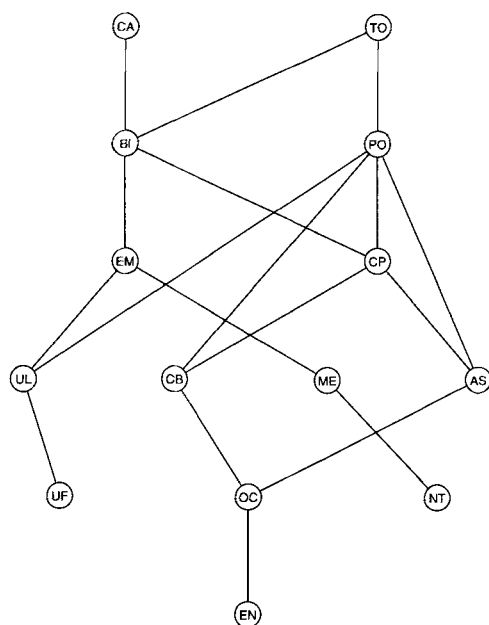
AN = Anzahl der Chemikalien

ID = Identifikationsmerkmale

IP = Informationsparameter

AR = Art der Chemikalien

QU = Qualität der Online-Datenbanken

Abb. 2: Hasse-Diagramm.
Scoring-Ansätze für 14 Online-Datenbanken

maximalen Elementen oder einfach von „Maximalen“. Diese beiden Objekte sind jedoch unvergleichbar untereinander. Zum Beispiel hat Chemical Abstracts eine höhere Bewertungsziffer für das Kriterium „Anzahl der Chemikalien“ als TOXLINE, jedoch für das Kriterium „Art der Chemikalien“ AR hat die Online-Datenbank TOXLINE ein höheres Score als Chemical Abstracts.

Die angesprochenen Maximalen sind, betrachtet man das Hasse-Diagramm, mit den Datenbanken vergleichbar, die

von jeweils einem der Maximalen ausgehend, durch einen Linienzug erreichbar sind. Dabei ist zu beachten, daß Linienzüge nach Regeln der Hasse-Diagrammtechnik nur von oben nach unten, oder (exklusiv) von unten nach oben gelesen werden dürfen. So ist im Hasse-Diagramm (\rightarrow Abb. 2) beispielsweise von TOXLINE aus POLLUTION, nicht aber COMPENDEX erreichbar.

Die Datenbank TOXLINE ist besser eingestuft als die Datenbanken ASFA, BIOSIS, CAB, EMBASE, ENVIROLINE, MEDLINE, NTIS, OCEAN, POLLUTION, UFORDAT und ULIDAT. Somit sind 11 Online-Datenbanken vergleichbar schlechter bewertet in bezug auf die oben genannten fünf Evaluationskriterien (Attribute in der Terminologie der Hasse-Algebra) als die Datenbank TOXLINE. Anhand des Diagramms kann man auch Folgen ablesen, in denen Datenbanken angeordnet sind, die in allen fünf Attributen gleichzeitig besser oder schlechter sind.

Beispiele hierfür sind:

EN < OC < CB < CP < BI < CA oder

UF < UL < PO < TO etc.

Diese beispielhaft gezeigten Anordnungsfolgen demonstrieren, daß eben keine strikte Ordnung mehr vorliegt, durch die alle Datenbanken gleichzeitig aufgereiht werden könnten. Vielmehr liegt nur noch eine partielle Ordnung vor, die Reihungen umfaßt, in denen nicht alle Datenbanken gleichzeitig auftreten.

Die Information, wieviele Datenbanken (allgemein Objekte) von einem Maximal erreichbar sind, ist wichtig genug, um in einer eigenen Kennzahl evident gemacht zu werden. Hierzu definieren wir für die Datenbanken die Menge „G“ als „Nachfolgermenge“ gesehen von oben nach unten und zählen die Anzahl ihrer Elemente.

So ergibt sich also beispielsweise für die Datenbank TOXLINE die Menge G wie folgt:

$G(\text{TO}) = \{\text{AS}, \text{BI}, \text{CB}, \text{EM}, \text{EN}, \text{ME}, \text{NT}, \text{OC}, \text{PO}, \text{UF}, \text{UL}\}$
 $\text{card}(\text{TO}) = 11$

Die Cardinalität einer Menge ist die Anzahl von Elementen der Menge, beschrieben durch die $\text{card}(G)$ für die Menge G.

Analog lassen sich die Mengen G und Cardinalitäten für die Online-Datenbank Chemical Abstracts festlegen:

$G(\text{CA}) = \{\text{AS}, \text{BI}, \text{CB}, \text{CP}, \text{EM}, \text{EN}, \text{ME}, \text{NT}, \text{OC}, \text{UF}, \text{UL}\}$
 $\text{card } G(\text{CA}) = 11$

Das bedeutet, daß beide Maximalen Chemical Abstracts und TOXLINE die gleiche Anzahl an Nachfolgermengen aufweisen.

Die Online-Datenbanken UFORDAT und ENVIROLINE sind beide minimale Elemente, „Minimale“, d.h. sie haben keine Nachbarn nach unten. Sie haben niedrige (schlechte) Bewertungsziffern für alle betrachteten Attribute.

Auffällig ist, daß bedeutende Umweltdatenbanken wie POLLUTION, ENVIROLINE und ULIDAT schlechter bewertet werden als die oben betrachteten Maximalen. POLLUTION befindet sich in der zweiten Ebene (von oben

nach unten gezählt) des Hasse-Diagramms, während ULDAT in der vierten Ebene und ENVIROLINE sogar in der sechsten Ebene des Hasse-Diagramms zu finden sind. Diese drei bekannten und bei vielen Hosts aufliegenden Umweltdatenbanken sind entweder bei dem Kriterium „Anzahl der Chemikalien“ oder bei den „Identifikationsparametern“ schlechter eingestuft als die diskutierten Maximale. Auffällig im Zusammenhang mit dem Bewertungskriterium „Identifikationsmerkmale“ ist die Tatsache, daß nur einige der ausgewählten Online-Datenbanken über die wichtigen Identifizierungsparameter, d.h. CAS-Nummern, Summenformel, Molekulargewicht verfügen. Keine der betrachteten Datenbanken verfügt über Strukturformeln. In dieser Beziehung sollten in Zukunft, unserer Meinung nach, die Datenbanken ergänzt und damit aufgewertet werden. Ebenso ist auffällig, daß die meisten der hier vorgestellten Online-Datenbanken ein hohes Score für das Kriterium „Informationsparameter“ (IP) erhalten haben. Das bedeutet, daß bibliographische umweltrelevante Online-Datenbanken eine Vielzahl von umweltrelevanten Parametern abdecken. Dieser Umstand trifft im Gegensatz hierzu nicht für alle numerischen Umweltdatenbanken zu.

5 Ausblick und weitere Forschungsarbeiten

Wir werden die Anwendung der Hasse-Diagramme in bezug auf die Einstufung und Bewertung von Datenquellen in Zukunft weiter ausbauen. Es wird dabei nicht nur an eine Ausweitung in bezug auf andere Datenquellen (z.B. CD-ROMs und Diskettenversionen von Datenbanken) und an andere Typen von Datenbanken (z.B. Faktendatenbanken) gedacht, sondern auch an einen verstärkten Einsatz von Methoden der diskreten Mathematik, z.B. der Graphentheorie. Eine tiefergreifende Betrachtung der Inhalte der Datenquellen im Zusammenhang mit der Thematik Chemikalien und Umweltschutz ist vorrangig. Publikationen hierzu sind in Vorbereitung.

Liste der verwendeten Akronyme und Abkürzungen

ASFA	Aquatic Science and Fisheries Abstracts
BIOSIS	Biosis Previews
BStMLU	Bayerischen Staatsministeriums für Landesentwicklung und Umweltfragen
CA	Chemical Abstracts
CAB Abstracts	Commonwealth Agriculture Bureaux
CIS	The Chemical Information System
COMPENDEX	Computerized Engineering Index
DIALOG	DIALOG Information Services, Inc.
DIMDI	Deutsches Institut für Medizinische Dokumentation und Information
ECDDIN	Environmental Chemicals Data Information Network
ECHO	European Host Organization
EMBASE	The Excerpta Medica Database
ENVIROLINE	Enviroline
ESA	European Space Agency
GSF	GSF-Forschungszentrum für Umwelt und Gesundheit GmbH
GBI	German Business Information
HSDB	Hazardous Chemicals Databank

ICSTI	International Centre for Scientific and Technical Information
MEDLINE	Medical Literature Analysis and Retrieval System Online
NTIS	National Technical Information Service
OCEAN	Oceanic Abstracts
POLLUTION	Pollution Abstracts
PUC	Projektgruppe Umweltgefährdungspotentiale von Chemikalien
STN	STN International The Scientific & Technical Information Network
TDS Numerica	Technical Database Services Inc.
TOXLINE	Toxicology Information Online
VINITI	The All Russian Institute of Scientific and Technical Information
UFORDAT	Umweltforschungsdatenbank
ULIDAT	Umweltliteratur-Datenbank

6 Literatur

- [1] YOUNG MARCACCIO, K.: Gale Directory of Databases 1994, Gale Research International Ltd., Andover (1994)
- [2] WILLIAMS, M.E.: Highlights of the Online/CD-ROM Database Industry: Opportunities Through Technology, Proceedings of the 14th National Online Meeting, New York, May 4-6, Medford NJ, Learned Information, Inc. (1993)
- [3] COX, J.: Environment Databases, An Aslib Online Guide, Aslib, The Association for Information Management, London (1991)
- [4] OTTAHAL, A.: Umwelt-Datenbank-Führer, Verlag TÜV Rheinland, Köln (1989)
- [5] VOIGT, K.; ROHLER, H.: Datenquellen für Umweltchemikalien, 2. Auflage, Ecomed Verlagsgesellschaft, Landsberg/Lech (1990)
- [6] VOIGT, K.: Gefahrstoff-Datenbanken, Leitung des Symposiums S4, Umweltbezogene Datenbanken Einführung, 9. Münchener Gefahrstoff-Tage 24.-27.11.93, Proceedings, 160-165, Ecomed Verlagsgesellschaft, Landsberg/Lech (1993)
- [7] VOIGT, K.: Informationsgewinnung aus Online-Datenbanken und CD-ROMs auf dem Gebiet der Naturwissenschaft und Technik, Vorlesung im Weiterbildungsstudiengang Umweltrecht & Umweltökonomie der Universität Lüneburg, Institut für Umweltwissenschaften am 24.10.93
- [8] Deutsches Institut für Medizinische Dokumentation und Information (DIMDI), Hostdefinierte Datenbankgruppen für Superbase und grips-Index, S. 1/13, grips-news Nr. 4/92
- [9] DIALOG Information Services, Inc., Database Catalogue, S. 150, DIALOG, Palo Alto (1994)
- [10] STN International, Datenbanken aus Wissenschaft und Technik, Broschüre, STN, Eggenstein-Leopoldshafen (1993)
- [11] VOIGT, K.; PEPPING T.: Informationssystem Umweltchemikalien, Abschlußbericht, GSF-Bericht 10/92, GSF-Forschungszentrum für Umwelt und Gesundheit GmbH, Neuherberg (1992)
- [12] VOIGT, K.; BENZ, J.; MÜCKE, W.: Datenquellen für Umweltchemikalien, Informationssystem Umweltchemikalien, Umweltwissenschaften und Schadstofforschung, 3/6 (1991) 367-369
- [13] BENZ, J.; VOIGT, K.; MÜCKE, W.: Strategy for Computer-Aided Searches for Information about Chemicals, Online Review, 13/2 (1989) 383-393
- [14] VOIGT, K.; PEPPING, T.: Datenbanken auf CD-ROM, in: HUTZINGER O., FIEDLER H., ECOINFORMA 92, 2nd International Congress and Exhibition on Environmental Information & Communication, Bayreuth 14.-18. September 1992, 247, ECOINFORMA Press, Bayreuth (1992)
- [15] MATTHIES, M.; ALTSCHUH, J.; BRÜGGEMANN, R.; VOIGT, K.: Data Needs, Data Availability and Data Estimation for Environmental Exposure and Hazard Assessment, in: STEINBERG C., KETTRUP A. (Eds.), Proceedings International Symposium on Ecotoxicology - Ecotoxicological Relevance of Test Methods, November 19-20, 1990, Neuherberg, GSF-Bericht 1/92, GSF-Forschungszentrum für Umwelt und Gesundheit GmbH, Neuherberg (1992)

- [16] VOIGT, K.; PEPPING, T.; KOTCHETOVA, E.; MÜCKE, W.: Testing of Online Databases in the Information System for Environmental Chemicals with a Testset of 68 Chemicals, *Chemosphere*, 24/7 (1992) 857–866
- [17] HARRY, V.; OPPENHEIM, C.: Evaluations of Electronic Databases, Part 1: Criteria for Testing CD-ROM Products, *Online & CD-ROM Review* 17/4 (1993) 211–222
- [18] HARRY, V.; OPPENHEIM, C.: Evaluations of Electronic Databases, Part 2: Testing CD-ROM Products, *Online & CD-ROM Review*, 17/4 (1993) 339–368
- [19] BRÜGGEMANN, R.; HALFON, E.: A Systematic Way to Analyse Hasse Diagrams, Theoretical Base of the Program „Hasse“, GSF-Forschungszentrum für Umwelt und Gesundheit GmbH, Neuherberg (1993)
- [20] HALFON, E.; REGGIANI, M.G.: On Ranking Chemicals for Environmental Hazard, *Environ. Sci. Technol.*, 20/11 (1986) 1173–1179
- [21] BRÜGGEMANN, R.; HALFON, E.: Ranking for Environmental Hazard of the Chemicals Spilled in the SANDOZ Accident in November 1986, *The Science of the Total Environment*, 97/98 (1990) 827–837
- [22] BRÜGGEMANN, R.; MÜNZER, B.: A Graph-Theoretical Tool for Priority Setting of Chemicals, *Chemosphere*, 27/9 (1993) 1729–1736
- [23] HALFON, E.: Comparison of an Index Function and a Vectorial Approach Method for Ranking Waste Disposal Sites, *Environ. Sci. Technol.*, 23/5 (1989) 600–609
- [24] HALFON, E.: Is There a Best Model Structure? *Ecological Modelling*, 20 (1983) 153–163
- [25] BRÜGGEMANN, R.; ALTSCHUH, J.: A Validation Study for the Estimation Aqueous Solubility from n-Octanol/water Partition Coefficients, *Science of the Total Environment*, 109/110 (1991) 41–57
- [26] BRÜGGEMANN, R.; MÜNZER, B.; HALFON, E.: An Algebraic/Graphical Tool to Compare Ecosystems with Respect to their Pollution – The German River „Elbe“ as an Example-I: Hasse-Diagrams, *Chemosphere*, 28/5 (1994) 863–872
- [27] BRÜGGEMANN, R.; STEINBERG, C.: Einsatz von Hasse-Diagrammtechnik zur vergleichenden Bewertung von aquatischen Wirkungstests, *Umweltwissenschaften und Schadstoff-Forschung*, akzeptiert 1995
- [28] VOIGT, K.; BRÜGGEMANN, R.: IsU, Informationssystem Umweltchemikalien, Information System for Environmental Chemicals – New Developments – in: Gesellschaft für Anlagen- und Reaktorsicherheit (GRS) mbH, WINRE '93, 4th Workshop on Information Management in Nuclear Safety, Radiation Protection, and Environmental Protection, 02.–04.11.93, Köln, GRS – 105, Proceedings, 179–188, Gesellschaft für Anlagen- und Reaktorsicherheit (GRS) mbH, Köln (1994)
- [29] VOIGT, K.; BRÜGGEMANN, R.: Metadatabases of Data-Sources for Environmental Chemicals, Proceedings 17th International Online-Information Meeting, 7–9 December 1993 London, S. 495–506, Learned Information Ltd., Oxford (1993)
- [30] VOIGT, K.; BRÜGGEMANN, R.: Ansatz eines Scoring-Systems für chemische/umweltrelevante Online-Datenbanken, Gesellschaft Deutscher Chemiker – GDCh, Fachgruppe Chemie-Information-Computer (CIC), 8. CIC-Workshop Jahrestagung 1993, Softwareentwicklung in der Chemie, 17.–19.11.93 Seeheim/Jugenheim, Proceedings, S. 1–15, GDCh, Frankfurt (1994)

Neues aus Forschung und Technologie

Multimedia

– Chance und Herausforderung

1 Einleitung

Der Erfolg einer modernen Telekommunikations- und Medienpolitik für Wirtschaft und Gesellschaft im 21. Jahrhundert hängt wesentlich von der inhaltlichen Gestaltung der neuen Multimedia-Technologien ab.

Schwierig gestaltet sich der Übergang zur Informationsgesellschaft durch Risiken, die sich aus den komplexen Eigenschaften digitaler Systeme und Restunzulänglichkeiten der Technik ergeben, z.B. Datensicherheit, Vertraulichkeit, mangelnde Nutzerfreundlichkeit, fehlender Schutz geistigen Eigentums oder eventuelle Softwarefehler. Nicht akzeptierbare **soziale Gefahren** bilden eine weitere ernstzunehmende Risikokategorie, wie die „Gläserne Gesellschaft“ und die „Zwei-Klassen-Gesellschaft“, ebenso wie Manipulation und Entfremdung der Anwender.

2 Chancen

Konventionelles Wirtschaftswachstum ist kein Garant mehr für künftigen Wohlstand. Die Grenzen bei Ressourcen, Umwelt, Verkehr und Transport sind in den hochindustrialisierten Ländern längst erreicht. Gleichzeitig vollzieht sich eine Globalisierung der Märkte, die Beschleunigung des technologischen Wandels und die Neuverteilung der Wirtschaftsstandorte mit einhergehender Neubestimmung der Standortmerkmale.

In der entstehenden Informationsgesellschaft verbinden breitbandige Kommunikationsnetze die verschiedenen Bereiche: öffentliche Verwaltungen, private Nutzer, Wirtschaft.

Alle Anwendungsfelder sind folgendermaßen technisch ausgestattet:

- ISDN
- TV-Kabelnetze
- breitbandige Glasfaser- bzw. Koaxialnetze
- Funknetze und Satelliten
- interaktive Endgeräte auf Anwenderseite

Multimediale Technik ist in folgenden Anwendungsgebieten vorstell- und machbar:

- **Medizin:** schnellere Hilfeleistung durch Ferndiagnosen und elektronische Patientenüberwachung
- **Soziale Dienste:** multimediale Betreuung älterer und behinderter Menschen
- **Wohnen:** individueller Konsumentenzugriff auf die Angebote des Weltmarktes durch Telebanking oder Teleshopping
- **Informations- und Unterhaltungsprogramme:** individuelle Mediennutzung durch „Video-on-Demand“ oder „Pay on View“
- **Öffentliche Verwaltungen:** mehr Bürgerfreundlichkeit durch elektronischen Zugriff (Autoanmeldung)