



„Kerngeschäft“ Sacherschließung in neuer Sicht

Was gezielte intellektuelle Arbeit und
maschinelle Verfahren gemeinsam
bewirken können


Dr. Gerhard Stumpf, UB Augsburg



Intellektuelle, „verstehende“ SE - notwendiger denn je, aber ...

- personalintensiv
- Web-Suchtechniken und Discovery statt proprietärer „bibliothekarischer“ Retrievalformen
- verbesserte automatische und maschinelle Verfahren
- Linked open data machen Output zum „Gemeingut“

Intellektuelle SE zurückgefahren – dennoch wachsen Effizienz und Reichweite



**Auch wenn die intellektuelle SE
als bibliothekarisches
Kerngeschäft zurücktritt, bleibt
sie für den Kern der
Metadatenwelt unverzichtbar.**

SE-Metadaten dienen...

in Katalogen und Retrievaltools

- zum primären Auffinden nach inhaltlichen Kriterien mittels Suchanfrage *[find]*
- als deskriptives Element bei der Anzeige *[identify]*
- zum sekundären Auffinden im iterativen Suchprozess ausgehend von angezeigten Metadaten *[collect]*
- zur Selektion nach inhaltlicher Relevanz für die jeweilige Fragestellung *[select]*

Funktion von Sacherschließungs-Metadaten bei der Suche am Beispiel von Schlagwörtern


Ihre Suchanfrage Thema/Schlagwort = logische partikel

Ausgewählte Suchfilter

Themen/Schlagwörter
(grob)

~~Deutsch~~

Treffer UB Augsburg (12)

1.  [Konditionalität und Kau...](#)
Volodina, Anna
2011

Ihre Suchanfrage Thema/Schlagwort = logische partikel

Treffer eingrenzen

Fach
[Germanistik / N...](#) (9)
[Sprach-/Literat...](#) (6)
[Slavistik](#) (1)

Themen/Schlagwörter
(grob)

[Logische Partikel](#) (12)

[Aufsatzsammlung](#) (3)

[Kausalsatz](#) (2)

[Syntax](#) (2)

[Textkohärenz](#) (2)

[Mehr anzeigen...](#)

Themen/Schlagwörter
(präzise)

[Adversativität ...](#) (1)
[Adversativsatz ...](#) (1)
[Deutsch / Adver...](#) (1)
[Deutsch / Appos...](#) (1)
[Deutsch / Aussa...](#) (1)
[Mehr anzeigen...](#)

Erscheinungsjahre

[>= 2006](#) (3)
[2002 - 2005](#) (3)
[1994 - 2001](#) (3)


UB Augsburg (9/12)

Substitutionsprozesse in Junktionen mit "dass" oder "ob"
Antworten auf Fragen zum Platzhalterproblem

VERFASSER: [Bischoff, Anneliese](#)
ORT, VERLAG, JAHR: Bielefeld, Bischoff, 1996
UMFANG: IV, 309 S. : graph. Darst.
SIGNATUR: 64/GC 7246 B621

SCHLAGWORT: [Deutsch](#) , [Fachsprache](#) , [Komplexer Satz](#) , [Logische Partikel](#) , [dass](#) , [ob <Wort>](#)

SCHLAGWORT 2: [Deutsch](#) , [Komplexer Satz](#) , [Konjunktion](#) , [Korrelat](#)

ISBN / ISSN: 3-925167-03-X 


[in die Merkliste](#)

Exemplare

[zum Dokument](#)

[mehr zum Titel](#)

Signatur	Standort	Zweigstelle	Status
Fußnoten			
01/GC 7246 B621	Zentralmagazin	UB	bestellbar


6.  [Handbuch der deutsche](#)
2003
64/GC 7117 P279

Grundleistungen der SE

Inhalte / Themen suchbar machen	Inhalte / Themen beschreiben	Themen, Entitäten und Ressourcen verlinken
über hinterlegte Indizes, maschinenlesbar	menschenlesbar und -verständlich	im Web, über Identifikatoren, Datenmodelle wie RDF
Metadaten: <ul style="list-style-type: none"> • Notationen und Benennungen (Klass.) • Bevorzugte und abweichende Namen • Codes 	Metadaten: <ul style="list-style-type: none"> • Benennungen (Klass.) • Bevorzugte Namen • Schlagwortfolgen • Abstracts 	Normdaten GND-interne Beziehungen GND – Wikipedia GND – VIAF usw.
Teile von Paratext und Volltext	Paratext Catalogue Enrichment	

Sacherschließung bezogen auf die FRBR-Entitäten der Gruppe I

Werk	Expression	Manifestation („Dokument“)	Exemplar
Inhaltsanalyse (erfolgt an Hand von Manifestationen)		Paratext (von Autor / Hrsg. / Verlag) <ul style="list-style-type: none"> • Inhaltsverzeichnis • Klappentext • Register • Abstract 	
<ul style="list-style-type: none"> • Zuweisung von Schlagwörtern für den Inhalt (Themen) • Klassifizieren des Inhalts 	Formschlagwörter für die Art des Inhalts Klassifizieren mit formal definierten Klassen (z. B. RVK-Notationen LP-LY)	Formschlagwörter für den Datenträger oder Materialtyp	<i>[Systematische Aufstellung gemäß Klassifikation]</i>
		Klassische Kataloganreicherung	
		Volltext-Extraktion	<i>[Provenienzerschließung (auch mit Normdaten)]</i>



Intellektuelle und maschinelle Sacherschließung

Intellektuelle Sacherschließung → Werk

Automatische Sacherschließung → Manifestation

Intellektuelle SE: Wörter → Begriffe → Themen

Automatische SE (Text-Extraktion): nur Wörter

Zusätzlich: probabilistische Verfahren

Ergebnisse aller Verfahren sollen interoperabel gemacht werden → Maschinelle Verfahren sollen möglichst ähnliche Ergebnisse bringen wie intellektuelle



Anreicherung von Titeldaten auf der Basis anderswo vorhandener Metadaten

- Optimierte Fremddatenübernahme
- Werkcluster-Verfahren: Kopieren von SE-Daten auf Titelsätze, die das gleiche Werk betreffen
- Anwendung von Crosskonkordanzen (Klassifikationen)
- Modifikation von SE-Daten, z. B. automatische Übersetzung
- Probabilistische Generierung von SE-Daten eines Systems aus Daten eines anderen Systems (z. B. fehlende Notationen aus vorhandenen Schlagwörtern)






Warum Anreicherungen mit z.T. unpräzisen Ergebnissen?

- Bibliothekarische Sacherschließung muss quantitativ relevant werden
- Kleine Inkonsistenzen müssen in Kauf genommen werden
- Konsistenz in Teilbereichen sollte bewahrt werden bzw. wiederherstellbar bleiben

Die Dichotomie von intellektuell vs. automatisch löst sich langsam auf.

Intellektuelle SE	Automatische SE
„Verstehende“ Inhaltsanalyse ist in der Lage, nicht adäquaten Paratext bzw. Metadaten zu erkennen	Linguistische Datenverarbeitung, z. B. Lemmatisierung, wörterbuchbasiert Grundlage: Volltext, Paratext, vorhandene Metadaten
Erzeugt Indexate in Anzeigeform (z. B. Schlagwortfolgen) mit dem Kriterium Verständlichkeit für Menschen	Erzeugt vor allem zusätzliche Sucheinstiege, nicht primär lesbare Inhaltsbeschreibung
Unabhängig von der Dokumentsprache, insofern konsistent	Immer abhängig von der Dokumentsprache
Kann tendenziell Unklarheit erkennen und Eindeutigkeit herstellen	Kann Verstehbarkeit im lesbaren Kontext nicht erkennen
Kann implizite Aspekte (z. B. geographische und sprachliche Bezüge) benennen	Wichtige, aber nicht explizit genannte Aspekte werden unterschlagen
Qualitativ überlegen	Qualitativ unterlegen
Quantitativ begrenzt	Großes quantitatives Potenzial

Die Dichotomie von intellektuell vs. automatisch löst sich langsam auf.

Intellektuelle SE	Maschinelle Verfahren zur Heterogenitätsbehandlung, Anreicherungen, Mappings, Crosskonkordanzen, Generierung von Normdaten, Visualisierung → Intellektueller Input wird z. T. deutlich aufgewertet	Automatische SE
		
		
	Alle Arten von SE-Metadaten können in großen Erschließungs- und Suchräumen interoperabel werden	

Ist Sacherschließung ohne Nutzung und Erzeugung von Metadaten möglich?

golfkrieg 1990 umweltschaden

Web Bilder Videos Shopping **Bücher** Mehr ▾ Suchoptionen

Ungefähr 182 Ergebnisse (0,59 Sekunden)

Tankerunfälle und der Ersatz ökologischer Schäden: am ..



<https://books.google.de/books?isbn=3825895564>

Liliane C. Kappet - 2006 - Vorschau

62 Von einem „Umweltschaden“ spricht Art. 2 Ziff. I Richtlinie 2004/35/EG zur Umwelthaftung: „Umweltschaden“ ist: (a) Schädigung geschützter Arten ... Die Waffenstillstandsresolution 687 vom 3.4.1991 nach dem **Golfkrieg 1990/91** (UN Doc.

Der Golfkrieg (1990/91): eine Auswahlbibliographie - Seite 10



<https://books.google.de/books?isbn=3922852440>

Ingeborg Otto, Marianne Schmidt-Dumont - 1992 - Snippet-Ansicht - Mehr Ausgaben

Englisch Persischer Golf + Irak-Kuwait-Krise / -Krieg (2. August 1990 - 11. April 1991) + Verlauf internationaler Konflikte + Umweltzerstörung + Umweltkatastrophe + Folgen von **Umweltschäden** + Umweltverschmutzung + Kriegsfolgen ...

Der Deutsche Wirtschaftsanwalt 2008/2009: Handbuch Fur ...



<https://books.google.de/books?isbn=3939804029>

Ein Handbuch für Unternehmen, Lexxion - 2008 - Vollansicht

... bei der Sanierung von **Umweltschäden**, die durch den ersten **Golfkrieg**

Snippets basieren i.d.R. nicht auf vorhandenen Metadaten, sondern auf Volltext, und stellen auch keine Metadaten dar.

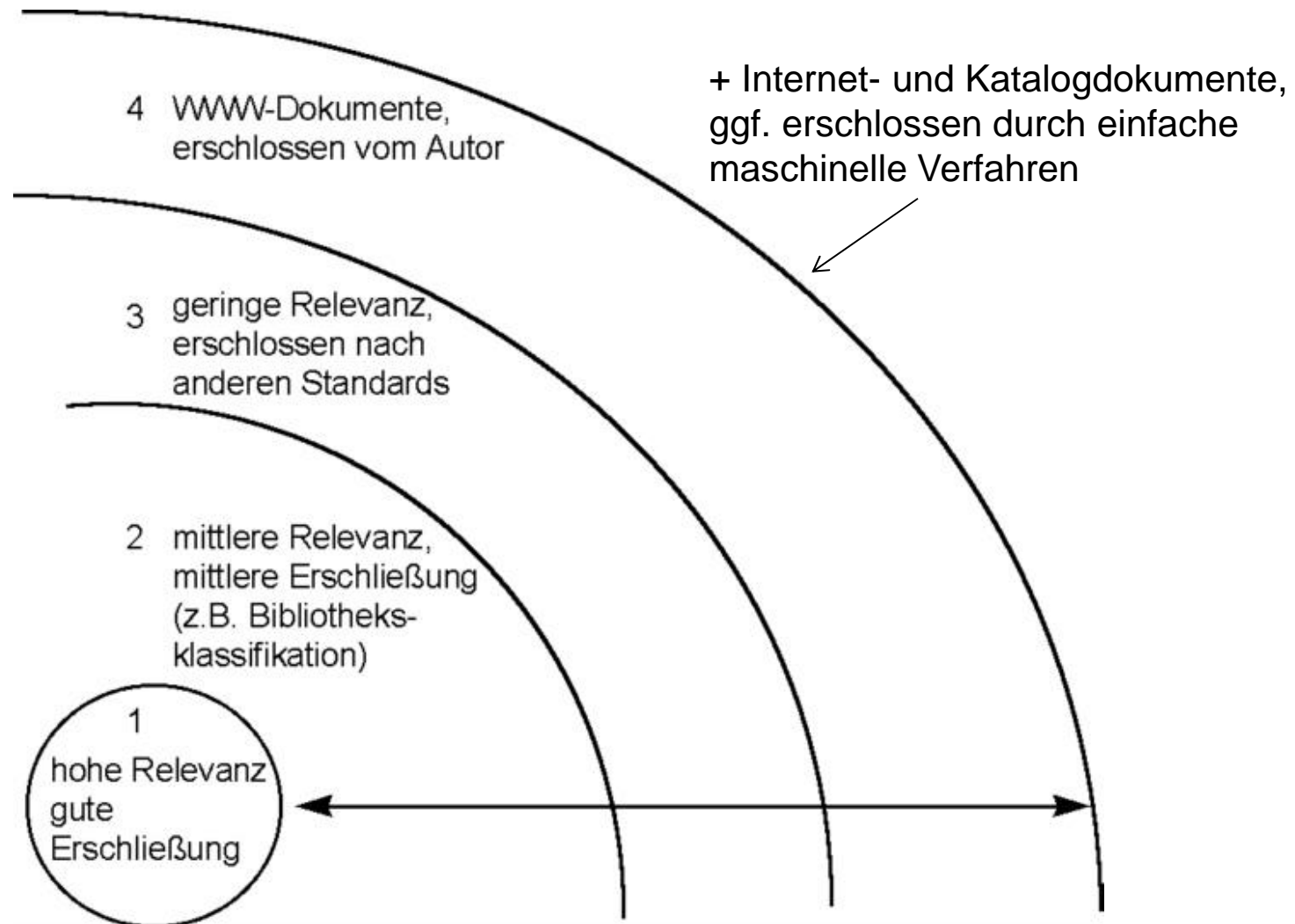


Neue Anforderungen an die Inhalterschließung

SE-Metadaten

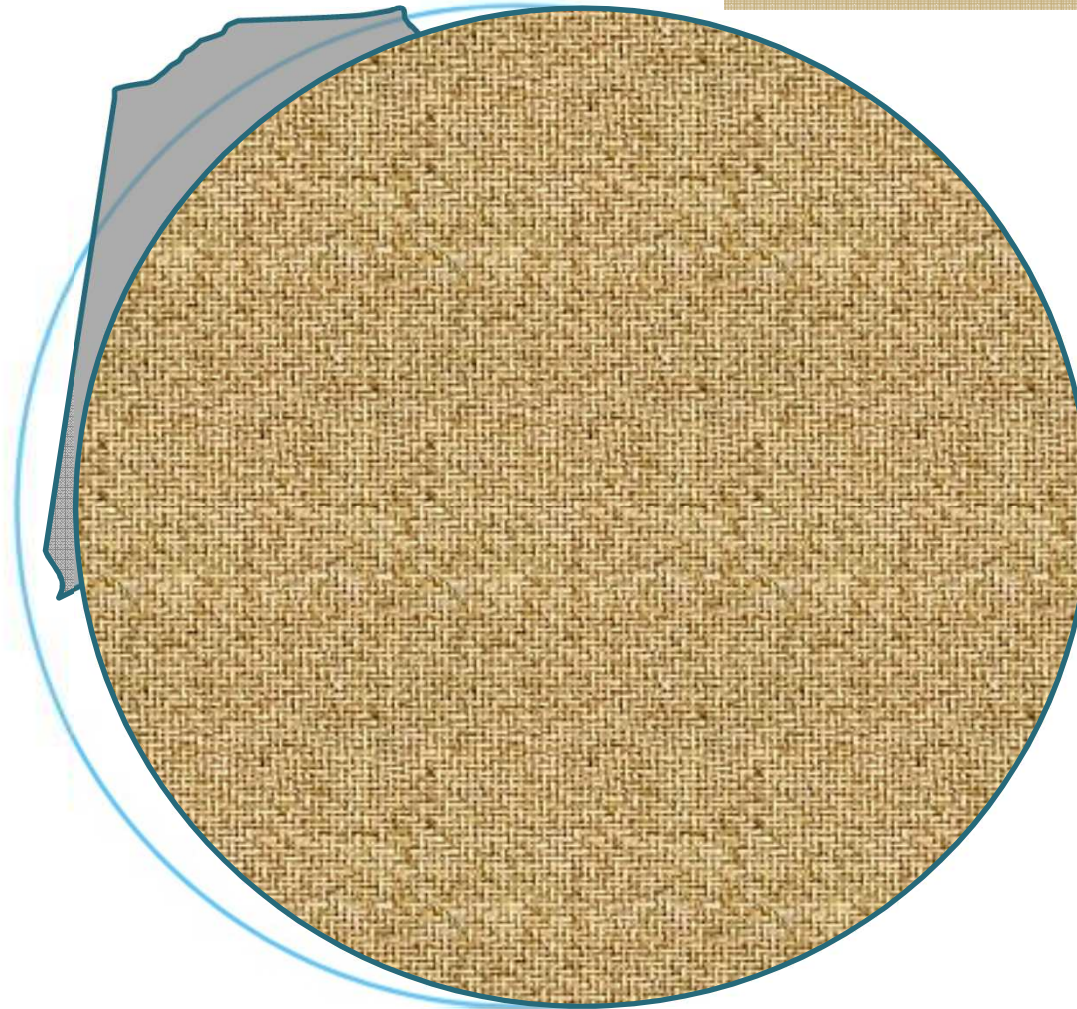
- dürfen nicht (mehr) nur auf einen fest definierten Suchraum abzielen,
- dürfen nicht nur mit bestimmten Suchtechniken nutzbar sein,
- sollen etablierten Standards folgen oder wenigstens mit solchen interoperabel gemacht werden,
- sollen direkt oder nach Weiterverarbeitung zur Nutzung als Linked data im Web geeignet sein,
- sollen neben Bibliotheken auch für andere Kulturbereiche attraktiv sein.


Schalenmodell nach J. Krause



Schalenmodell

Einzelne Erschließungsräume mit qualitativ heterogener Erschließung fließen in großen Erschließungs- und Suchräumen zusammen.






Kontrollierte Heterogenität nach dem Schalenmodell

Standards und Schemata aus dem Kernbereich bestimmen auch die Strukturen, auf die hin die Daten der äußeren Schalen bearbeitet und angereichert werden.

„... Raum für eine neue Sichtweise auf Konsistenzerhaltung und Interoperabilität zu schaffen.: Die Standardisierung ist von der verbleibenden Heterogenität her zu denken.“ (DFG-Positionspapier *Aktuelle Anforderungen der wissenschaftlichen Informationsversorgung*, Mai 2004)

Durch bessere Möglichkeiten des Datenmanagements wird Standardisierung im „klassischen“ Sinn nicht überflüssig, sondern wirksamer.



Neuausrichtung der Sacherschließung nach dem Leitbild der Discovery-Systeme?

- Inhaltlich weiterführende Facetten sind essenziell wichtig, Konsistenz bzw. Interoperabilität vor allem innerhalb der Facetten erforderlich.
- Die Differenzierung wissenschaftlicher Thematiken verlangt ausreichende Granularität der Erschließung.
- Die entitätsbezogene, ins Web geöffnete Verlinkung (d. h. die Nutzung von Normdaten) stellt die größte Herausforderung bei der Weiterentwicklung der inhaltlichen Discovery dar.
- Das Entdecken soll das Finden ergänzen und erweitern, nicht ersetzen. Präzise „Kernfragen“ müssen mit hochwertiger Erschließung aus dem Kern beantwortet werden.

Mehrwert der intellektuellen Erschließung

- Kontextualisierung

Syntagmatisch-sequenzieller Kontext	Kontext im semantischen Netz
Fügt Einzelelemente (nicht zwingend Normdaten) zu einer Themenbeschreibung zusammen (präkombiniert oder präkoordiniert)	Führt von Einzelelementen der Inhaltsbeschreibung (Normdaten) in ein semantisches Umfeld, erschließt weitere Themen und Ressourcen
Sprachlich-begrifflich orientiert; Grundmuster: menschenlesbarer Text	an Entitäten orientiert; Grundmuster: maschinell auswertbares Geflecht von Aussagen über Ressourcen

- Syntagmatische Kontextualisierung kann - über eine vage „Nützlichkeit“ von Wörtern hinaus - Kernthemen hervorheben, Einzelaspekte gewichten, kognitiv eingängige Strukturen schaffen.

Mehrwert der intellektuellen Erschließung

1	Golfkrieg 1990-1991 Türkei Kurden Flucht Geschichte 1991 Jugendsachbuch
2	Golfkrieg 1990-1991 Umweltbelastung
1	Golfkrieg 1990-1991 Umwelthaftung Schadensersatzanspruch Auswirkung Internationales Umweltrecht Völkerrecht
2	Golfkrieg 1990-1991 Umweltkatastrophe
7	Golfkrieg 1990-1991 Umweltschaden
1	Golfkrieg 1990-1991 Umweltschaden Aufsatzsammlung
1	Golfkrieg 1990-1991 Umweltschaden Korallenriff Persischer Golf
1	Golfkrieg 1990-1991 Umweltschaden Ölunfall Meeresverschmutzung Aufsatzsammlung
1	Golfkrieg 1990-1991 Umweltverschmutzung
4	Golfkrieg 1990-1991 United Nations

#/Treffer	Autor	Titel
1/14	Hawley, T. M.	Against the fires of hell
2/14		The environmental aftermath of the Gulf War
3/14	Golob, Richard S.	Environmental impact of war in the Arabian Gulf
4/14	Farīd, Ḥātim N.	Ġarīma d-eidda al-ḥeayāt
5/14	Ṣādiq, Muḥammad	The Gulf War aftermath
6/14	Bāz, Fārūq al-	The Gulf War and the environment
7/14	Payne, Cymie R.	Gulf War reparations and the UN Compensation Commission
8/14	Bloom, Saul	Hidden casualties
9/14	Vogt, Helge Peter	Korallenriffe im Persischen Golf
10/14	Linden, Olof	Kriegets miljøeffekter
11/14	Kühnel, Matthias	Ökologische Auswirkungen des Golfkrieges
12/14	Sartorius, Rolf [Bearb.]	Ökologische Auswirkungen des Golfkrieges
13/14	Mackinnon, Michael	Quellen der Finsternis
14/14	*Abū-Zināda, *Abd-al-*Azīz H.	The status of coastal and marine habitats two years after the Gulf War oil spill

Beispiel: thematische Umfeldsuche mit Schlagwortfolgen: liefert relevante Treffer in allen Sprachen

Mehrwert der intellektuellen Erschließung

golfkrieg 1990 umweltverschmutzung

Web Bilder Videos News Shopping Mehr ▾ Suchoptionen

Ungefähr 45.200 Ergebnisse (0,54 Sekunden)

Zweiter Golfkrieg – Wikipedia
de.wikipedia.org/wiki/Zweiter_Golfkrieg ▾
Dieser Artikel behandelt den Krieg in Irak und Kuwait zwischen 1990 und 1991. ... des Kriegsendes und den hohen Grad an mittelbaren **Umweltschäden** aus.

Infoblatt Zweiter Golfkrieg - Ernst Klett Verlag - Schulbücher ...
www2.klett.de/sixcms/list.php?page=geo...Golfkrieg...Golfkriege ▾
Der Zweite **Golfkrieg** erstreckt sich über zwei Phasen: Die Krise begann mit dem Einmarsch ... Der **Golfkrieg** 1990/91 in Irak und Kuwait und das Kurdenproblem ... Eine der größten **Umweltkatastrophen** der Geschichte ereignete sich auf dem ...

[PDF] Der zweite Golfkrieg (1990-1991) - KOBRA - Universität ...
https://kobra.bibliothek.uni-kassel.de/bitstream/urn.../Mansur_Diss1.pdf
12.05.2006 - 1.0 Stand der Forschung zum Thema zweiter **Golfkrieg** (1990-1991)
Hinsicht, ist es die Fähigkeit nicht von der **Umwelt** lernen zu müssen.

12 Jahre nach dem Golfkrieg, 17.01.2003 (Friedensratschlag)
www.ag-friedensforschung.de/regionen/irak/12-jahre-golfkrieg.html ▾
17.01.2003 - Die Gesundheits- und **Umweltkosten** eines Krieges gegen den Irak ... Er untersucht die Auswirkungen des **Golfkrieges** von 1990 bis 1991 auf ...

Umwelt & Ökologie - Amazon
www.amazon.de/...Golfkrieg-Umwelt.../s?...n%3A189467%2CK%3AZw... ▾

Vergleichbare
Suchmaschinensuche in
Webressourcen




Intellektuelle Arbeit gezielt einsetzen

Prioritäten setzen (Definition des Kernbereichs für die eigene Institution):

- Welche Wirkung erzielt die Erschließungsleistung im gesamten potenziellen Wissens- und Suchraum?
- Wie relevant ist die Erschließungsleistung für die eigene Community / Haupt-Nutzerschaft?
- Welchen Mehrwert verspricht die Erschließungsleistung im Kontext anderer Erschließungsmöglichkeiten?

Kann intellektuelle Sacherschließungskompetenz eingesetzt werden

- zum Ausbau des semantischen Netzes über die Pflege von Normdaten,
- zur Erstellung von Vorgaben für maschinelle Verfahren (z B. Crosskonkordanzen),
- für Test, Training und Evaluation automatischer Erschließungsmodule?



**Herzlichen Dank für Ihre
Aufmerksamkeit und für alle
weiter führenden Fragen und
Überlegungen!**