

# Suchstrategien im Deep Web für SportwissenschaftlerInnen

AGSB – Jahrestagung  
Salzburg

Florian Seiffert  
12.04.2006

# Übersicht

- Deep Web
- Salzburger These
- Deep Web vs. Google
- SportPilot
- Suchstrategien Sportwissenschaft
- Tips für eigene Suchstrategien

# Deep Web

- Unter 'Deep Web' oder 'Hidden Web' wird der Teil des World Wide Web verstanden, der via Suchmaschinen nicht erreichbar ist.
- gemeint sind:
  - Datenbanken,
  - Bibliothekskataloge,
  - aus Content-Management-Systemen erzeugte Seiten (Zeitungen, Zeitschriften, ...)

# Bloody Deep Web

- Unter 'Bloody Deep Web' wollen wir den Teil des World Wide Web verstehen, der weder über Suchmaschinen noch sonstwie erreichbar ist, der aber erreichbar sein sollte!

# Bloody Deep Web

z.B.:





# Ultra optimistische Vision



**Jules Verne**  
*Von der Erde  
zum Mond*

*Roman · Diogenes*

# Salzburger These

- Bis zum Jahr 2031 sind **alle** sportwissenschaftlich relevanten Informationen im 'World Wide Web' erreichbar!
  - Metadaten
  - Volltexte
    - Bücher, Zeitschriftenartikel, Aufsätze, ...
  - Fernseh- und Rundfunksendungen
    - Bundesliga, Olympische Spiele, WM, ...
  - Datenbanken

# Deep Web vs. Google

- Google und Co:
  - Eine Suchzeile
  - „dahinter“ der Inhalt von ca. 8 Milliarden Seiten
  - Niemand muß sich Details merken:
    - Woher kommt der Inhalt
    - Wo finde ich genau, was sich Suche
    - Wie komme ich zu richtigen Seite
    - Tipp-Fehler werden verschmerzt
- Deep Web:
  - Für jede Datenbank brauche ich
    - Details und Spezialkenntnisse
    - evtl. eigene Suchsyntax
    - Kenntnisse einer neuen Oberfläche



# Deep Web - Google

- Google und Co:
  - Suchmaschinen sind **EINFACH**
  - Suchmaschinen werden **häufig benutzt**
- Deep Web:
  - Datenbanken und Kataloge machen **MÜHE**
  - Datenbanken werden **seltener benutzt**
- Problem(e):
  - Viele Schätze bleiben verborgen / werden nicht gefunden
  - ExpertInnen werden gebraucht, die „Laien“ weiterhelfen

# Deep Web - Google

- Wie kann man das Problem lösen?
  - Strategie 1
    - Schulungen
    - Kurse
    - Einführungen in Literaturrecherche
    - ...
    - **eine Google-artige Sportsuche bauen**
  - Strategie 2
    - Das Deep Web sichtbar machen, damit Google es findet.
    - **Für jedes Buch der ZBSport eine eigene HTML-Seite**

# google-artige Sportsuche: SportPilot / ViFa-Sport

- Möglichst alle Sportdatenbanken einbinden
- Dem Suchenden die Entscheidung abnehmen: „Wo muß ich suchen“
- Eine Suchzeile (zunächst)
- Schnelle Antwortzeiten

The screenshot shows the website 'Zentralbibliothek der Sportwissenschaften' of the 'Deutsche Sporthochschule Köln'. The page features a navigation menu on the left with options like 'STARTSEITE', 'ADRESSE', 'KONTAKT', 'ÖFFNUNGSZEITEN', 'KATALOG', 'SUCHE', 'DATENBANKEN', 'ELEKTRONISCHE ZEITSCHRIFTEN', 'BESTÄNDE', 'DISSERTATIONEN', 'AUSLEIHE', 'FERNLEIHE', and 'TUTORIUM'. The main content area displays the title 'Zentralbibliothek der Sportwissenschaften' and a list of bullet points describing the library's services, including being the largest special library for sports science in Germany and a central scientific library for the German Republic. Below the text is a search bar labeled 'SportPilot:' with a 'Suche' button and a help icon. A message 'Hier (noch) keine KATALOG-Suche!' is displayed above the search bar. The footer contains the text 'Web-Redaktion der ZBSport, 05.04.2006'.

# google-artige Sportsuche: SportPilot / ViFa-Sport

- Eingebunden:
  - Web-Seiten ZBSport
  - Pressemeldungen DSHS
  - Gießener Alterssportarchiv
  - Sportpersonen (Munzinger)
  - CurrentContent
- Geplant:
  - Katalog der ZBSport
  - Sportdatenbanken (ViFa-Sport)
  - Sportmedizinische Literatur (ZBMed)
  - aktuelle Sportmeldungen (sid, dpa, kicker, ...)
  - Einfach „alles zum Sport“



# für jedes Buch eine eigene HTML-Seite:

- Virtueller Nachbau der Bücherregale der ZBSport
- Jede Seite enthält alle Informationen zu dem Buch mit
  - Signatur (Standort)
  - Links zu Bestellmöglichkeit
- statisches HTML, damit Suchroboter (Google und Co) die Seiten indexieren
- Bestand der ZBSport wird in Google findbar

# für jedes Buch eine eigene HTML-Seite:

- Anreicherung jeder Seite mit allen kodierten Informationen:
  - Systematik der ZBSport (5 Fus 100a = Zeitschriften Fußball)
  - Systematik der ZDB
  - Schlagwörter
  - Normdateien (Verweisungsformen)
- Auflösung von Abkürzungen
  - Zeitschriftentitel (Z. , J. ...)
- „aufwärmen“ - „erwärmen“
- Rebounder – Trampolin
- Fußball - Soccer

# Suchstrategien SportPilot

## Beispiele 1:

• Karlsruher Motorik Screening	0
• Sportunterricht 2004	5
• Bös	13
• Trainingspläne	1
• Trainingspläne Trainingskarten	0
• Trainingspläne	1
• Bertke	0
• kurzausleihe	0
• ausleihdauer	0
• Nikephoros	19
• Thorpe	2

# Suchstrategien SportPilot

## Beispiele 2:

• reitkunst	0
• borowik	0
• athletik jahrbuch	0
• Kosten	16
• Kosten überzogen	0
• tritt drauf, der zuckt noch	0
• castellani	1
• murmelspiele	0
• tourist frontier of Nouveau Quebec	0
• hochschulschriften	80
• klein-vogelbach	30



# Suchstrategien Fazit

- Es wird „einfach“ gesucht
  - Das Trunkierungszeichen (\*) wird kaum eingesetzt
  - Das Ausschließungszeichen (-) wird garnicht eingesetzt
  - Wenig Mehrwortsuchen
- Suche zu ca. 80% nach **Titel** und/oder **AutorIn**
  - Offenbar wird meist nach bereits bekanntem gesucht

# Suchstrategien Fazit

- Es wird offenbar angenommen, dass die Maschine versteht, was ich meine:
  - „neueste Forschungsergebnisse im Sport“
  - „Inhalte neuester Sportstudien“
  - „wie fühlt der fan sich im stadion?“
  - „warum gehen die fans ins stadion?“

# Suchstrategien Folgerungen

- Wenn man eine „Suche“ baut:
  - Mach es möglichst einfach
    - Groß-Klein-Schreibung, Zeichensatz, wenig Suchfelder, wenig Datenbankfelder
  - Die Suche ist für andere, nicht für Dich
- Wenn man selbst sucht:
  - Maschinen sind „blöd“
    - Variieren Sie Ihre Suchworte
      - Singular, Plural, deutsche Fälle, Englisch, Originalsprachlich
      - Diss., Dissertation, Doktorarbeit
    - Suchen Sie konkret, keine Oberbegriffe
      - Köln statt Karnevalshochburg am Rhein
      - London statt Hauptstadt Großbritannien

# Kontakt: Florian Seiffert

Zentralbibliothek der Sportwissenschaften Köln  
[Seiffert@DSHS-Koeln.de](mailto:Seiffert@DSHS-Koeln.de)

Ich bedanke mich für die Aufmerksamkeit!

## Fragen?