

# Online Recherche

Marcus Lindemann,

Recherchetrainer und  
geschäftsführender Autor, autoren(werk)

[lindemann@autorenwerk.de](mailto:lindemann@autorenwerk.de)

[www.autorenwerk.de](http://www.autorenwerk.de)

[www.recherche-info.de](http://www.recherche-info.de)



# Marcus Lindemann (41)

- Gründer und geschäftsführender Autor, autoren(werk)  
Produktionen u.a. für Frontal21, WISO
- Journalismusdozent und Trainer,  
Schwerpunkt Recherche, on- und offline sowie TV-Journalismus
  - Akademie für Publizistik
  - ARD/ZDF Medienakademie
  - Universität Leipzig
  - Universität Darmstadt
  - Konrad-Adenauer-Stiftung
  - Friedrich-Ebert-Stiftung
  - Evangelische Medienakademie und Journalistenschule
  - direkt für Sender, Regionalzeitungen, Agenturen
- Rechercheur, dann Reporter und Redakteur im ZDF
- freier Journalist für zahlreiche Regionalzeitungen, TEMPO, DeutschlandRadio, Mitarbeit an arte-Themenabenden des SWF
- MA in political economy
- Mitglied im netzwerk recherche (nr), Investigative Reporters and Editors (IRE)



# Googleisierung

*In dem Dorf Waffensen nahe dem niedersächsischen Rotenburg an der Wümme scheint die Welt noch in Ordnung: Der Gasthof "Eichenhof" lockt mit gemütlicher Kaminschenke, Bauer Poppe um die Ecke verkauft Fleisch und Marmelade aus eigener Produktion, und bislang brachte allein der Shanty-Chor einen Hauch der großen, weiten Welt in den norddeutschen Heideflecken.*

*(Spiegel)*



# Googeln Sie | Waffensen |

- Auf welchen Plätzen der Trefferliste finden Sie Bauer Poppe, den Eichenhof und den Shantychor?
- Was sagt das über die Arbeit des Spiegel?
- Wie und warum könnten sich der reale Ort Waffensen und das virtuelle Waffensen (rekonstruiert über Google) unterscheiden?



# Googleisierung

- **1. Google definiert heute den Mindeststandard journalistischer Recherche.**
- **2. Google enttarnt Schablonen-Journalismus und Fakes.**
- **3. Google definiert oft den Maximalstandard journalistischer Recherche.**
- **4. Google bietet ein verzerrtes Bild der medialen Welt. Dieses Zerrbild wirkt wiederum auf die Medien zurück.**

**(Jochen Wegner, 2005)**



# Im Internet steht so viel Falsches ...

- Bei Twitter wird soviel Mist erzählt.
- Blogs kann man auch nicht trauen.
- Die Wikipedia kann jeder manipulieren.
- Im Netz weiß man nie, mit wem man es zu tun hat.
- Einzelne Fehler im Netz beschädigen immer noch das ganze Medium.

# Vorsicht! Gefahr!



The image shows a screenshot of the German federal government website (bundesregierung.de) with a search result. The page features a navigation menu on the left with links for 'Startseite', 'Service', 'Suche', 'Ministerien', and 'Tipps'. The main content area displays a search result for 'BUNDESKANZLERIN MERKEL TRITT ZURÜCK'. The search bar contains 'bundesregierung.de' and 'Bundesministerien'. The headline reads 'BUNDESKANZLERIN MERKEL TRITT ZURÜCK'. The text below the headline states: 'Bundeskanzlerin Dr. Angela Merkel ist aus persönlichen Gründen mit sofortiger Wirkung von ihrem Amt als Bundesrepublik Deutschland zurückgetreten. Stoiber hat angekündigt die Führungsgewalt zu übernehmen zum Endsieg über den Terrorismus zu führen, notfalls mit Gewaltanwendung. In Kürze mehr.' At the bottom of the page, there is a copyright notice: '© 2006 Presse- und Informationsamt der Bundesregierung'.

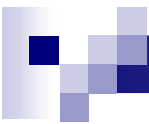


# Andererseits:

## Haben Sie schon mal ...

- ... verifiziert, wer Sie angerufen hat?
- ... verifiziert, wen Sie angerufen haben?
- ... überprüft, ob eine Pressemitteilung tatsächlich vom genannten Urheber stammt?
- ... geprüft, ob ein Fax vom tatsächlichen Absender stammt?
- ... Fehler in einem Lexikon gefunden?





# Warum gibt es Vorbehalte gegen Online-Quellen?

- Jeder kann veröffentlichen.
- Technische Vorbehalte/fehlendes Verständnis.
- Informationen verbreiten sich schnell.
- Fehlende Unterscheidung zwischen Inhalt und Medium.
- Es gibt tatsächlich viel Falsches und Halbwahres im Netz.



# Offline- vs. Online-Recherche

- Es gibt keine grundsätzlichen Unterschiede zwischen Online- und offline-Recherche.
- Entscheidend ist in beiden Sphären die Bewertung der Quellen.



# Quellen-Analogien

## offline

Amtliche Bekanntmachungen  
(Bundesanzeiger, Amtsblätter,  
Gesetzestexte, Handelsregister)

Archive, Bibliotheken,  
Datenbanken

Dokumente, Geschäftsberichte,  
Pressemitteilungen von Unter-  
nehmen, Vereinen, Behörden etc.

Insider-Informationen

Aussagen von Akteuren,  
Betroffenen, Experten

Stammtische, Gerüchte

## online

identisch, wenn sie von amtlichen  
Sites stammen

weitgehend identisch (Vorteil online:  
Volltext durchsuchbar!)

Homepages der entsprechenden  
Institutionen

Versehentlich online-gestellte  
Informationen (Bsp:| “not for  
publication“ site:.mil |), Wikileaks

identisch (wenn aus guten Quellen)

Groups, Foren etc.



# Vorteile der Online-Recherche

- 24/7 – Recherche rund um die Uhr weltweit
- kostengünstig
- unbemerkte, auch verdeckte Recherche möglich (nicht automatisch gegeben!)
- kann, muss aber nicht schneller sein als offline Recherche



# Achtung: Online ist nicht alles

## Nachteile der Online-Recherche:

- Oft zeitaufwändig, verführt zur Recherche in die Breite.
- Informationen, die nicht online sind, werden ausgeblendet.
- Online ist kein Abbild von offline (siehe Googleisierung)



# Effizientes Recherchieren

- Wechsel zwischen Online- und Offline-Quellen (mal Recherche über das Telefon schneller, mal über das Netz)
- Sehr selten kann man ein Thema nur offline oder nur online recherchieren
- Selbst wer fast alles offline recherchiert, kann online noch Zeit sparen.



# Bewertung von Online-Quellen.

## Beispiele:

- BBC-Interview zum 20 Jahrestag der Chemie-Katastrophe von Bhopal

<http://www.youtube.com/watch?v=LiWlvBro9eI>

BBC hat Dow-Sprecher über diese Domain angefragt:  
dowethics.com

- martinlutherking.org
- bunddeutscherjuristen.org

(Übung: Was ist auf diesen Seiten jeweils auffällig, wer steckt hinter den Seiten?)



# Welchen Online-Quellen kann ich trauen?

- Privilegierten Quellen wie Websites von Behörden, Ministerien, bundestag.de, handelsregister.de
- Primärquellen  
Im Zweifelsfall so nah an die Quelle wie möglich.
- Quellen, die mit Hilfe einer zweiten (unabhängigen) Quelle verifiziert werden können.





# Wie kann ich Online-Quellen überprüfen?

- 1.) Inhalt der Seite, erster Eindruck, Rechtschreibung, Stil, Design etc. Objektivität, Interessenlage, Aktualität
- 2.) Anbieter ähnlicher Informationen suchen, Inhalte vergleichen
- 3.) Wer verlinkt auf die Seite?
  - `link:www.domainname.de`
- 4.) Bewertungen der Site durch Dritte, ggf Namenssuche (inkl. Datenbanken)
- 5.) Impressum auswerten,
  - ggf. Impressum mit Registrierungsdaten vergleichen.
  - ggf Namenssuche (inkl. Datenbanken)
- 6.) besser: Registrierungsdaten: whois bzw Denic.
- 7.) Alte Internetseiten (archive.org)
- 8.) Fortgeschritten & teuer: Domain History, weitere Domains des Betreibers.



# Und was ist mit Wikipedia?

- Verzicht wäre töricht.
- Deshalb:

## **Nutzen ja, abschreiben nein.**

- Nutzen für die Recherche:
  - Oft direkte Links zu guten Quellen/Primärquellen.
  - Finden besserer Suchbegriffe, möglicher Quellen.
  - Einzelne Fakten aus der Wikipedia können gegooglet werden – auf der Suche nach zweiten Quellen



# Wie funktioniert eine Suchmaschine?



# Wie funktioniert eine Suchmaschine?

- 1.) Suchmaschinen durchsuchen das Netz systematisch und speichern die gefundenen Daten auf Ihren eigenen Rechnern in den sogenannten Index.



# Wie funktioniert eine Suchmaschine?

- 1.) Suchmaschinen durchsuchen das Netz und erstellen einen Index.
- 2.) Die eingegebenen Suchbegriffe werden in diesem Index gesucht, d.h. ***nicht live im Internet!***

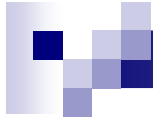


# Wie funktioniert eine Suchmaschine?

- 1.) Suchmaschinen durchsuchen das Netz und erstellen einen Index.
- 2.) Die eingegebenen Suchbegriffe werden in diesem Index gesucht, d.h. **nicht live im Internet!**
- 3.) Die gefundenen Seiten werden sortiert nach Bedeutung und Beliebtheit sortiert (Kriterien weitgehend bekannt: Verlinkung, Häufigkeit der Suchbegriffe, Pagerank etc.)



# Was Google alles nicht findet...



# Was google alles nicht findet...

1. Inhalte, die nach dem letzten Besuch des google-robots online gestellt wurden

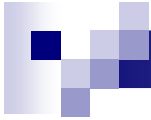




Das Internet ist sooooo aktuell...

Beispiel:

<http://recherche-info.de/handouts-recherche-kann-man-lernen/uebungen-und-loesungen/suchseite/>



# Was google alles nicht findet...

1. Inhalte, die nach dem letzten Besuch des google-robots online gestellt wurden
2. Google indexiert nicht alle Websites und auch nicht alle erfassten Websites werden vollständig indexiert.



# Google zensiert

- Pornografie
- Rechtsradikales Gedankengut
- „Inhalte, die gegen nationale Rechtsprechung verstoßen“

Images Showing: All image sizes Results 1 - 20 of about 21,900 for tiananmen. (0.09 seconds)

Tip: Save time by hitting the return key instead of clicking on "search"



**Tiananmen**  
705 x 742 pixels - 141k - jpg  
jml.prof.free.fr



Remember **Tiananmen Square**  
640 x 403 pixels - 58k - jpg  
www.loc.gov



the advance of a tank column with ...  
1124 x 794 pixels - 231k - jpg  
www.christusrex.org



**Tiananmen.jpg**  
1092 x 616 pixels - 309k - jpg  
www.freedomtocare.org



... **Tiananmen 20Square 202.jpg**  
494 x 449 pixels - 23k - jpg  
www.rollins.edu



Remember **Tiananmen**  
640 x 452 pixels - 53k - jpg  
jamesandannie.cyberflunk.com



**tiananmen**  
256 x 256 pixels - 16k - gif  
www.guardian.co.uk



**Tiananmen Square Protest**  
264 x 400 pixels - 31k - jpg  
dreaming\_big\_dreams.blogspot.com



tiananmen

402 tiananmen 1-20 0.13



19981002\_TONY-TIANANMEN.jpg 897 x 601 - 147k - jpg sp.cs.tsinghua.edu.cn



Mrs. Gutierrez at Tiananmen Square 1728 x 1152 - 317k - jpg www.usembassy-china.org.cn



Tiananmen Square Night Scenery 500 x 348 - 21k - jpg www.silkroute.cn



Tiananmen Square 500 x 375 - 170k - jpg www.silkroute.cn [ www.silkroute.cn ]



tiananmen.jpg 255 x 191 - 9k - jpg grid.hust.edu.cn



tiananmen.jpg 500 x 350 - 14k - jpg www.pku.edu.cn



chinatravel-tiananmen 200 x 144 - 10k - jpg www.china-tour.cn



tiananmen.jpg 200 x 143 - 18k - jpg eaie.njtu.edu.cn



TianAnMen.jpg 1468 x 954 - 79k - jpg db.pku.edu.cn



Click to see the big black & white ... 600 x 408 - 67k - jpg www.csist.org.cn



# Was google alles nicht findet...

1. Inhalte, die nach dem letzten Besuch des google-robots online gestellt wurden
2. Google indexiert nicht alle Websites und auch nicht alle erfassten Websites werden vollständig indexiert.
3. Zensierte Inhalte (Pornografie, Rechtsextremismus, diverses)



# Was google alles nicht findet...

1. Inhalte, die nach dem letzten Besuch des google-robots online gestellt wurden
2. Google indexiert nicht alle Websites und auch nicht alle erfassten Websites werden vollständig indexiert.
3. Zensierte Inhalte (Pornografie, Rechtsextremismus, diverses)
4. Inhalte, von denen der domain-Betreiber nicht will, das sie gefunden werden.



# robots.txt

## Beispiele:

- [whitehouse.gov/robots.txt](http://whitehouse.gov/robots.txt)
- [basf.com/robots.txt](http://basf.com/robots.txt)





Suchkategorien

Produkte & Preise

News

Service & Hilfe

Kontakt



- Firmeninformationen
- Bundesanzeiger
- Tages- und Wochenpresse
- Fachpresse
- Branchenberichte & Fakten
- Personeninformationen
- Wissenschaftliche Nachweise

Quellenangebot

- Quellenauswahl
- Direkt zur Quelle:

DB-Name oder Kürzel

mit Übernahme der Suche

GO >>

Mein GENIOS

Neuregistrierung >

KENNUNG

\*\*\*\*\*

Subaccount (optional)

Login >>

Passwort vergessen?

Sie wollen sich Dokumente oder Artikel ansehen, ausdrucken, abspeichern?

Zugang als registrierter Profikunde

Vorhandene Kennung eingeben oder als Profikunde registrieren (Ihre Vorteile: Mengenrabatt, AboService, Persönliche Einstellungen uvm.)

Kennung  Zur Neuregistrierung  
Passwort  Passwort vergessen?  
Subaccount (optional)  **OK >>**

Einmaliger Zugang als Gastkunde über

Kreditkarte  
(VISA, MasterCard, American Express)



Bitte beachten Sie, dass ein **Mindestrechnungsbetrag in Höhe von 4,76 EUR** anfällt. (Preise inkl. MwSt.)

ClickandBuy  
per Bankeinzug oder Kreditkarte  
(Was ist ClickandBuy?)



T-Pay  
(MicroMoney, Telekom-Rechnung, Lastschrift)



Hier können Sie unsere **Allgemeinen Geschäftsbedingungen** einsehen.



# Was google alles nicht findet...

1. Inhalte, die nach dem letzten Besuch des google-robots online gestellt wurden
2. Google indexiert nicht alle Websites und auch nicht alle erfassten Websites werden vollständig indexiert.
3. Zensierte Inhalte (Pornografie, Rechtsextremismus, diverses)
4. Inhalte, von denen der domain-Betreiber nicht will, das sie gefunden werden (robots.txt)
5. Passwortgeschützte Inhalte
6. Inhalte von Datenbanken werden nur teilweise gefunden
7. Interaktiv generierte Seiten
8. (Zum Teil:) Bezahlinhalte



# Was google alles nicht findet...

1. Inhalte, die nach dem letzten Besuch des google-robots online gestellt wurden
2. Google indexiert nicht alle Websites und auch nicht alle erfassten Websites werden vollständig indexiert.
3. Zensierte Inhalte (Pornografie, Rechtsextremismus, diverses)
4. Inhalte, von denen der domain-Betreiber nicht will, das sie gefunden werden (robots.txt)
5. Passwortgeschützte Inhalte
6. Inhalte von Datenbanken werden nur teilweise gefunden
7. Interaktiv generierte Seiten
8. (Zum Teil:) Bezahlinhalte Google indexiert nicht alle Websites und auch nicht alle erfassten Websites werden vollständig indexiert.
9. Seiten, die „nicht mehr online sind“, die aber online archiviert sind.



# Was google alles nicht findet...

1. Inhalte, die nach dem letzten Besuch des google-robots online gestellt wurden
2. Google indexiert nicht alle Websites und auch nicht alle erfassten Websites werden vollständig indexiert.
3. Zensierte Inhalte (Pornografie, Rechtsextremismus, diverses)
4. Inhalte, von denen der domain-Betreiber nicht will, das sie gefunden werden (robots.txt)
5. Passwortgeschützte Inhalte
6. Inhalte von Datenbanken werden nur teilweise gefunden
7. Interaktiv generierte Seiten
8. (Zum Teil:) Bezahlinhalte Google indexiert nicht alle Websites und auch nicht alle erfassten Websites werden vollständig indexiert.
9. Seiten, die „nicht mehr online sind“, die aber online archiviert sind.
10. Texte in Bildern, Videos und Texte in anderen Dateiformaten, die Google nicht lesen kann.



# Was heißt das für die Recherche?

- Google ist nicht das Internet.
- Es gibt Millionen von Seiten, die kostenfrei zugänglich sind, aber nicht von Google gefunden werden.
- Häufiger Quellen suchen, nicht Inhalte.



# Nutzung professioneller Datenbanken

- Handelsregister
- Auskunftsteien (z.B.: Creditreform)
- Pressearchive
- Fachspezifische Datenbanken



# Nutzung professioneller Datenbanken

- Liste von Unternehmen, die zwischen 1933 und 1945 gegründet wurden
- Volltextsuche im eigenen Zeitschriftenregal
- Was bringt die Lockerung des Kündigungsschutzes?  
Liste aller Unternehmen mit weniger als 20 Mitarbeitern für eine Region/Postleitzahl
- Liste von Unternehmen, die in den Iran oder den Irak exportieren



# Erweiterte Suche von Google

Wichtige Syntax:

„“, OR, -

filetype, site, intitle/inurl etc.

Sprache & Region: meist keine gute  
Einschränkung

Datum: Achtung liefert auch zahlreiche fehlerhafte  
Ergebnisse!

related/link





# Syntax außerhalb der erweiterten Suche

- \*
- +
- ~
- Numrange:
- Phonebook
- define:
- info:



# Andere nützliche Dienste von Google

- Bilder- und Video-Suche (mit je eigener erw. Suche)
- News (hier: Sehr gute Datumssuche!)
- Bücher/Blogs
- Groups: gute Suche in Foren und Usenet
- Alerts
- Translate: zumindest zur Übersetzung der Navigation nützlich
- Scholar: Suche in wiss. Quellen
- Reader: RSS-Feed-Reader



# Derzeit unverlinkte Google-Services für Journalisten

- Google-Trends
  - Google- Verzeichnis
  - benutzdefinierte Suchmaschinen
  - Google public data explorer
- 
- Nicht mehr online:  
Gaudi (google audio indexing)



# Google macht mehr...


... als Sie wollen

- Reihenfolge: Paris Hilton vs. Hilton Paris
- Suchdomain: google.com und google.de
- Ort des Internetzugangs: Suchergebnisse außerdem Standort-abhängig
- Schreibweisen: *Photographie* findet auch *Fotografie* (unterbinden durch „“ oder +)
- Automatische Trunkierung  
z.B.: Suchbegriff *journalistischen*  
findet auch *Journalismus* und *journalistische*



# 10 Gebote für erfolgreiche Internet-Recherchen

1. Fragen Sie sich, **was Sie suchen!**  
(Welche Ergebnisse erwarten Sie von Ihrer Recherche?)  
*Nutzen Sie Formulierungen/Begriffe, die Sie im Ergebnis erwarten.*
2. Benennen Sie die **möglichen Quellen** für diese Informationen!  
(Deep Web Quellen finden Sie oft auf dem Umweg über potenzielle Urheber.)
3. Wählen Sie die **angemessenen Recherchemittel!**  
(Eignet sich das Internet überhaupt? In welcher Phase der Recherche eignet es sich? Vergessen Sie Telefon und Oeckl nicht!)
4. Vergessen Sie nichts von dem, was Sie sonst beim Recherchieren beachten würden!  
(**Bewertung der Quellen**, Interessen der Quellen, zweite Quelle, Möglichkeit von abbeschriebenen Fehlern...)
5. **Lesen Sie** die Ergebnisse, die Sie finden, während der Recherche!  
*Verfeinern Sie damit Ihre Suchanfrage!*

- 
6. *Nutzen Sie mehrere Suchbegriffe, bevorzugen Sie Substantive gegenüber Verben, bringen Sie die Suchbegriffe in die Reihenfolge Ihrer Wichtigkeit.*
  7. Nutzen Sie bei umfassenden Recherchen immer *mehrere Recherchewege* im Internet!  
(verschiedene **Suchmaschinen, Kataloge, Institutionen, Usenet, deep web**)
  8. Wenn Sie nichts finden, fragen Sie sich, **warum Sie nichts finden.**  
(*Suchen Sie bei neuen Datenbanken, Suchmaschinen etc. immer zunächst nach Informationen, von denen Sie wissen, dass sie da sind.*)
  9. **Pflegen Sie Ihre Lesezeichen** wie Ihre Informanten!  
(In Ordner sortieren, aktualisieren)
  10. Bleiben Sie auf dem Laufenden, nicht nur in Ihrem Ressort, sondern verfolgen Sie auch die Fortentwicklung von Suchmaschinen etc.



# Minimalstandards für die Internet-Recherche

- Kenntnis der Syntax und Funktionsweise von Suchmaschinen.
- Versuchen die Zahl der Treffer zu reduzieren, Ergebnisse einzugrenzen
- Methodik statt Fleißarbeit, denn Fleiß statt Methode führt zu Fehlern
- Systematisches Abarbeiten von Trefferlisten
- Auffinden von Seiten, die Suchmaschinen (noch) nicht finden können.
- Finden der eigenen Suchbegriffe auf den Trefferseiten (Strg F)
- Einsatz der site-spezifischen Suche
- Recherchefragen erkennen, die offline gelöst werden müssen
- effiziente Kombination von On- und Offline-Recherche
- Bewertung von Online-Quellen



# Noch Fragen?

Mailen Sie mir,  
wenn Sie mit einer Recherche nicht  
weiterkommen!

[lindemann@autorenwerk.de](mailto:lindemann@autorenwerk.de)





Das war's!  
Danke!

[www.recherche-info.de](http://www.recherche-info.de)