

# „Die Ente bleibt draußen!“

**Workshop über Quellenprüfung online  
Mit ein paar Anmerkungen  
und Merksätzen zur Online-Recherche**

**Albrecht Ude**

Bibliothekarstag  
Mi., 23. Mai 2012



## Um mich Ihnen vorzustellen:

Albrecht Ude – Journalist / Rechercheur / Recherche-Trainer

<http://www.ude.de/> .

<http://eine-woche-ohne.de/> .

<http://recherche-standards.wikispaces.org/> .

<http://www.netzwerkrecherche.de/> .



## Was man online so alles recherchieren kann:

Am Beispiel von Bildern

1. Die Bildinhalte (allgemeiner: Dateiinhalte)
2. Die Metadaten [→ Beispiel: albrecht.jpg]
3. Die Internet-Adresse (URL)  
[http://a2.twimg.com/profile\\_images/1492236703/chris-high-res-cropped\\_normal.png](http://a2.twimg.com/profile_images/1492236703/chris-high-res-cropped_normal.png) .

Wichtig: Searchable Information - maschinell recherchierbare Daten



## Online recherchieren: Bilder – Inhalte

Manipulierte Bilder eines iranischen Raketentests (2008)

<http://www.heise.de/tp/artikel/28/28295/1.html> .

... und die Reaktion im Netz:

[http://larrinski.files.wordpress.com/2008/07/2658316482\\_32be3e6f21.jpg](http://larrinski.files.wordpress.com/2008/07/2658316482_32be3e6f21.jpg) .



## Online recherchieren: Bilder – Adressen

Vorüberlegung:

So ein Winzbild ist maschinell erzeugt, kein Mensch würde ein so kleines Bild herstellen - Menschen stellen größere Bilder von sich ins Netz (ungefähr Passbildgröße)

Machinell erzeugt bedeutet, dass auch der Dateiname automatisch entstanden ist. Es muss irgend ein Muster im Dateinamen geben.

### 1.) Bildsuche nach dem kleinem Bild

[http://a2.twimg.com/profile\\_images/1492236703/chris-high-res-cropped\\_normal.png](http://a2.twimg.com/profile_images/1492236703/chris-high-res-cropped_normal.png) .

Die Google-Bildsuche liefert weitere Bilder, Links und den Namen der Person



## Online recherchieren: Bilder – Adressen

### 2.) Ermitteln des großen (ursprünglichen) Bildes

[http://a2.twimg.com/profile\\_images/1492236703/chris-high-res-cropped\\_normal.png](http://a2.twimg.com/profile_images/1492236703/chris-high-res-cropped_normal.png) .

Durch etwas Probieren kann man aus der Adresse des kleinen Bildes auf die des ursprünglichen Bildes schließen - meist muss man dazu Teile der URL löschen oder ändern:

[http://a2.twimg.com/profile\\_images/1492236703/chris-high-res-cropped.png](http://a2.twimg.com/profile_images/1492236703/chris-high-res-cropped.png) .

Die Bildsuche nach dem großem Bild liefert die besten Ergebnisse, vor allem Hinweise auf weitere Links zur Person.



## Online recherchieren: Bilder – Adressen

### 3.) Suche nach dem Dateinamen

[http://a2.twimg.com/profile\\_images/1492236703/chris-high-res-cropped\\_normal.png](http://a2.twimg.com/profile_images/1492236703/chris-high-res-cropped_normal.png) .

Selbstverständlich kann man bei jeder Suchmaschine nach dem Dateinamen suchen. Das geht auch mit dem ermittelten Namen der großen Bilddatei



## Online recherchieren: Bilder – Adressen

### 4.) Ermitteln des Servers (twimg -> twitter, also site:twitter.com)

[http://a2.twimg.com/profile\\_images/1492236703/chris-high-res-cropped\\_normal.png](http://a2.twimg.com/profile_images/1492236703/chris-high-res-cropped_normal.png) .

Über eine Domainrecherche wird klar, dass "[a2.twimg.com](http://a2.twimg.com)" zu Twitter gehört. Folglich kann man Recherchen mit "[site:twimg.com](http://a2.twimg.com) OR [site:twitter.com](http://a2.twimg.com)" schärfen





## Online recherchieren: Bilder – Adressen

### 5.) Recherche nach der Nummer

Die Nummernfolge "1492236703" in der URL (ein Verzeichnisname) ist mit hoher Wahrscheinlichkeit eine Accountnummer oder Postingnummer [http://a2.twimg.com/profile\\_images/1492236703/chris-high-res-cropped\\_normal.png](http://a2.twimg.com/profile_images/1492236703/chris-high-res-cropped_normal.png).

Die Suche danach mit dem Operator "[site:twitter.com](#)" liefert ein Posting, dass auf die Profilseite bei Twitter verlinkt. Ebenso kann man mit dem Operator "[inurl:1492236703](#)" arbeiten.



# 1. Erfolgreiche Recherchen beginnen im Kopf / Google ist ein Umweg

Beispiel: Suche nach Volker Hauth  
(Anfrage eines Kollegen einer Dänischen Tageszeitung per E-Mail im August 2011)

Dear Albrecht,

My good colleague [XXX] said that you might be able to help me: **I am looking for a phone number or email for Volker Hauth; a German architect/engineer who studied at the Technical University of Hamburg-Harburg together with Mohamed Atta - the ringleader from the 11. september-attack.** I can't find a number or info for Hauth on the internet, so I was hoping that you might know where to look.  
Hope to hear from you and thanks.



# 1. Erfolgreiche Recherchen beginnen im Kopf / Google ist ein Umweg

Google nach "volker hauth" – Ungefähr 29.400 Ergebnisse (0,14 Sekunden)

Viel zu viel, um ökonomisch zu recherchieren.

- ✓ Wer Architektur studiert, wird vermutlich Architekt.
- ✓ Wer in Hamburg studiert, bleibt möglicherweise dort.
- ✓ Woher kriege ich Informationen über Architekten in Hamburg?

Architektenkammer Hamburg

<http://www.akhh.de/index.php?id=33> .



## 2. Google ist eine scharfe Waffe (wenn man damit umgehen kann)

Die Google Operatoren: + - OR "" \* .. site: filetype: .

Wer diese Operatoren beherrscht (es gibt noch ein paar mehr), kann schon viel.

Beispielhaft ist die schönste Online-Recherche von 2008: "Hack the Olympics" von Strydehax:

Wie alt ist He Kexin?

[http://de.wikipedia.org/wiki/He\\_Kexin](http://de.wikipedia.org/wiki/He_Kexin) ,

<http://strydehax.blogspot.com/2008/08/hack-olympics.html> .



### 3. Wikipedia ist keine Quelle

Karl Theodor zu Guttenberg in der "Bild" und anderen Medien als "Wilhelm" bezeichnet, weil jemand diesen flaschen Vornamen in die Wikipedia eingeflutscht hatte:

<http://www.bildblog.de/5704/wie-ich-freiherr-von-guttenberg-zu-wilhelm-machte/> .

<http://www.bildblog.de/5731/wilhelm-ii/> .

[http://de.wikipedia.org/wiki/Karl-Theodor\\_zu\\_Guttenberg](http://de.wikipedia.org/wiki/Karl-Theodor_zu_Guttenberg) .

[http://de.wikipedia.org/w/index.php?title=Karl-Theodor\\_Freiherr\\_von\\_und\\_zu\\_Guttenberg&diff=56419545&oldid=56418683](http://de.wikipedia.org/w/index.php?title=Karl-Theodor_Freiherr_von_und_zu_Guttenberg&diff=56419545&oldid=56418683) .



### 3. Wikipedia ist keine Quelle

Wenn Sie Wikipedia nutzen (nicht nur lesen) wollen, achten Sie nicht nur auf die Artikel, sondern checken Sie immer:

Versionsgeschichte – Diskussion – Versionsgeschichte der Diskussion



## 4. Wikipedia hilft bei Navigation und Kontextualisierung

Bei aller Skepsis gegenüber den Artikeln - Wikipedia ist ein grossartiges Werkzeug für die Navigation und zum Einordnen und Erweitern von Wissen!

Wikipedia als Sprungbrett zu autoritativen Quellen:

Weblinks: Literatur von und über Karl-Theodor zu Guttenberg im Katalog der Deutschen Nationalbibliothek -> DNB PND

<https://portal.dnb.de/opac.htm?query=Woe%3D129602434> .

Normdaten (ganz unten auf der Seite): Karl-Theodor zu Guttenberg in der Personennormdatei:

<http://d-nb.info/gnd/129602434> .



## 4. Wikipedia hilft bei Navigation und Kontextualisierung

Web-Katalog: Ausgewählte, von Menschen (!) für wichtig befundene Websites:

[http://de.wikipedia.org/wiki/Karl-Theodor\\_zu\\_Guttenberg#Weblinks](http://de.wikipedia.org/wiki/Karl-Theodor_zu_Guttenberg#Weblinks) .

Presse- / Quellensammlung zum Thema: Einzelnachweise

[http://de.wikipedia.org/wiki/Karl-Theodor\\_zu\\_Guttenberg#Einzelnachweise](http://de.wikipedia.org/wiki/Karl-Theodor_zu_Guttenberg#Einzelnachweise) .





## 4. Wikipedia hilft bei Navigation und Kontextualisierung

Veröffentlichungen von und Literatur über zu Guttenberg:

[http://de.wikipedia.org/wiki/Karl-Theodor\\_zu\\_Guttenberg#Ver.C3.B6ffentlichungen\\_.28Auswahl.29](http://de.wikipedia.org/wiki/Karl-Theodor_zu_Guttenberg#Ver.C3.B6ffentlichungen_.28Auswahl.29) ,

[http://de.wikipedia.org/wiki/Karl-Theodor\\_zu\\_Guttenberg#Literatur](http://de.wikipedia.org/wiki/Karl-Theodor_zu_Guttenberg#Literatur) .

Verweis:

[http://de.wikipedia.org/wiki/Plagiatsaff%C3%A4re\\_Guttenberg](http://de.wikipedia.org/wiki/Plagiatsaff%C3%A4re_Guttenberg) .



## 5. "Die Ente bleibt draußen" – prüfen Sie Ihre Quellen

Der Bund Deutscher Juristen

<http://www.bunddeutscherjuristen.org/> .

Domainsuche für de-Domains

<http://www.denic.de/> .

Domainsuche für alle Domains

<http://www.iana.org/> – Root Zone Database

Der Domaincheck - wem gehört eine Domain? - ist eine Pflichtrecherche!



## Werbung

Firefox! Firefox! Firefox!

Für tiefe, analytische Recherchen gibt es keinen besseren Webbrowser als Firefox.

- ✓ Open-Source ;
- ✓ Schnelle Reaktion auf Sicherheitslücken ;
- ✓ ca. 2.000 Zusatzprogramme (Add-ons, Erweiterungen) .

Bestimmte Formen der Recherche sind nur mit dem Firefox möglich!



## Ausgangslage

- ✓ Das Internet / WWW schafft technische Voraussetzungen für Meinungsfreiheit.
- ✓ Meinungsfreiheit ist nicht Wahrheit / Richtigkeit.
- ✓ Das Web wimmelt von veralteten / falschen / gefälschten Informationen.
- ✓ Wer kommuniziert / publiziert, tut dies interessengeleitet.
  
- ✓ US-Neonazis publizieren über Martin Luther King,
- ✓ Deutsche Bundesrichter fordern die Folter,
- ✓ Nachrichten über Raffineriebrände lassen Ölpreise steigen,
- ✓ Parteien "bearbeiten" Wikipedia-Einträge politischer Konkurrenten,
- ✓ ... a never ending story!



## Interne Kriterien

### Sichtbare Informationen der Webseite

Diese Kriterien können sie auch nutzen, wenn Sie nur einen Ausdruck der Webseite haben.

#### Autor

- ✓ genannt? [wenn nein, → Impressum → Quellcode]
- ✓ bekannt? (für einen bestimmten Standpunkt / Interesse / Vorurteil?)
- ✓ anerkannte Autorität?
- ✓ hat er eine eigene Homepage?
- ✓ was weiß Google\* über ihn?



## Interne Kriterien Sichtbare Informationen der Webseite

Diese Kriterien können sie auch nutzen, wenn Sie nur einen Ausdruck der Webseite haben.

### Autor

- ✓ genannt? [wenn nein, → Impressum → Quellcode]
- ✓ bekannt? (für einen bestimmten Standpunkt / Interesse / Vorurteil?)
- ✓ anerkannte Autorität?
- ✓ hat er eine eigene Homepage?
- ✓ was weiß Google\* über ihn?

\* nutzen Sie Google nicht ...

... allein: Nutzen Sie auch Ask, Bing, DuckDuckGo, Exalead, Yahoo, und Yandex

... oder die Metasuchmaschinen Dogpile, IxQuick und Metager2.



## Interne Kriterien Sichtbare Informationen der Webseite

### Publikationsdatum

- ✓ genannt? [→ Quellcode, → HTTP Header]
- ✓ korrekt? [→ Caches]
- ✓ aktuell oder veraltet? (Wenn alt: gab es Überarbeitungen?)

Die gleichen Fragen gelten für das Datum von Aktualisierungen.



## **Interne Kriterien**

### **Sichtbare Informationen der Webseite**

#### **Die Website – Urheber / Herausgeber / Medium** (gleiche Fragen wie zum Autor)

- ✓ was für eine Site ist das –  
wissenschaftlich, journalistisch, privat, Wiki?
- ✓ wer veröffentlicht dort noch?
- ✓ wo ist der Standort? (lokale Gesetze / Zensur)





## Interne Kriterien

### Sichtbare Informationen der Webseite

#### Der Text

- ✓ googlen\* Sie nach dem Titel
- ✓ wo ist der Text erwähnt / zitiert?  
(googlen\* Sie nach URL, Titel, eindeutigen Sätzen)
- ✓ von wem?
- ✓ wie ist der Text dort bewertet? (bewerten Sie Erwähnungen!)
- ✓ falls nirgends erwähnt – woran könnte das liegen?
  
- ✓ Original oder Plagiat?



## Interne Kriterien

### **Sichtbare Informationen der Webseite**

### **Inhalt(e) des Textes**

- ✓ worum geht's?
- ✓ für wen ist der Text geschrieben?
- ✓ ist die Argumentation plausibel?
- ✓ ist die Information stabil?
- ✓ sind die Zitate korrekt?
- ✓ welcher Standpunkt / welches Vorurteil wird vertreten?
- ✓ können Sie die Fakten verifizieren? Wie?
- ✓ vergleichen Sie den Text mit anderen Quellen.
- ✓ ist der Text ironisch, satirisch oder ein Hoax?



## Interne Kriterien Sichtbare Informationen der Webseite

### Sonderfall Wikipedia

- ✓ achten Sie auf die "History" (frühere Versionen) und wichtige Änderungen des Artikels,
- ✓ achten Sie auf die Diskussionen der Bearbeiter
- ✓ tobt gerade ein "edit war"?
- ✓ wer bearbeitet diesen Artikel? Wie viele Bearbeiter? Was bearbeiten sie noch?

Nutzen Sie Wiki-Watch!



## Interne Kriterien Sichtbare Informationen der Webseite

### Die Webseite als Datei

- ✓ ist der Autor des Textes auch Urheber der Seite?
- ✓ ist die Seite stabil, oder wird sie verschwinden / sich ändern?
- ✓ ist das (technische) Format angemessen?
- ✓ welche Soft- und Hardware brauchen sie, um die Seite zu sehen?



## Externe Kriterien

### Die Webseite als Teil der Website

#### Die Website

- ✓ gibt es ein Impressum?
- ✓ enthält es korrekte Kontaktdaten?
- ✓ wie finanziert sich das Ganze?
- ✓ gibt es Informationen über Sponsoren?
- ✓ was ist deren Interesse?



## Externe Kriterien

### Die Webseite als Teil der Website

#### Der URL (Uniform Resource Locator, “Web-Adresse”)

- ✓ passt der URL zum Inhalt?  
(**.gov** : US Government; **.de** : Deutschland; **.com** : Kommerziell ...)
- ✓ was für ein URL: Uni, Bibliothek, Firma, Weblog, Medium, anders?
- ✓ Ist eine Session-ID Teil des URL?
- ✓ Ist der URL stabil?

( Wenn Sie die Seite zitieren wollen,

- ✓ speichern Sie sie ,
- ✓ schauen Sie, ob Sie sie bei [archive.org](http://archive.org) oder in stabilen Caches finden )



## Externe Kriterien

### Die Webseite im Netz

#### Externe Links

- ✓ wer verlinkt auf diese Seite? [ → Google "link:" ]
- ✓ wie viele externe Links gibt es?
- ✓ wie wird die Seite dort erwähnt?
- ✓ gibt es Links von "Authorities" (z.B. DMOZ,)?
- ✓ falls es keine externen Links auf die Seite gibt, woran könnte das liegen?
- ✓ finden Sie die Seite bei Bing, Google, Yahoo, ...?
- ✓ falls nein, woher kennen Sie *diese* Seite?



## Externe Kriterien

### Die Webseite im Netz

#### Links auf externe Seiten

- ✓ gibt es Links auf externe Seiten?
- ✓ falls nein, woran könnte das liegen?
- ✓ wohin zeigen die Links?
- ✓ funktionieren die Links? [ → Firefox-Erweiterung "LinkChecker" ]
- ✓ wie wichtig / abseitig sind die verlinkten Seiten?





## Externe Kriterien

### Die Webseite als Datei

### Der Quellcode

Immer in den Quellcode gucken! [ Ansicht -> Seitenquelltext anzeigen ]

- ✓ achten Sie auf Kommentare  
(zwischen `<!--` und `-->` oder `//-->` )
- ✓ lesen Sie den HTML-Header:  
Infos über Autor, Urheber, Daten usw.
- ✓ wenn nötig, testen Sie den Quellcode.



## Externe Kriterien (2)

### Versteckte Daten und Metadaten

### Directory stemming

- ✓ Versuchen Sie "directory stemming"
- ✓ Nutzen Sie Google\* dafür (**+site:domain.tld +”index of”**)

Anm. für Fortgeschrittene:

- ✓ Beschäftigen Sie sich mit den Standards von Webserver-Software (am verbreitetsten ist Apache)
- ✓ Lernen Sie Fehlermeldungen provozieren und lesen



## Externe Kriterien (2)

### Versteckte Daten und Metadaten

### Eingebundene Dateien

- ✓ Gibt es eingebundene Dateien (Bilder, Downloads wie PDF, DOC, XLS ...)?

Fast jeder Dateityp hat eigene Metadaten - prüfen Sie diese!



## Externe Kriterien (2)

### Versteckte Daten und Metadaten

#### Robots.txt

- ✓ Gibt es einen "Robots Exclusion File"  
( [www.domain.tld/robots.txt](http://www.domain.tld/robots.txt) )
- ✓ Welche Dateien / Verzeichnisse sollen Suchmaschinen ignorieren?



## Externe Kriterien (2)

### Versteckte Daten und Metadaten

### Whois

Jede Domain ist in einer Whois-Datenbank des zuständigen Network Information Center (NIC) erfasst.

Sie finden alle NICs in der Root Zone Database der IANA (Internet Assigned Numbers Authority).

- ✓ was steht im Whois-Eintrag?
- ✓ ist die Information komplett und valide?
- ✓ falls nein, woran könnte das liegen?



## Externe Kriterien (2)

### Versteckte Daten und Metadaten

### Web-Caches

Sie finden viele Caches im Netz, z.B. bei Google und Yahoo, das Coral Content Distribution Network oder Archive.org.

Archive.org, die “WayBackMachine” archiviert Websites.

- ✓ Suchen Sie nach der Seite in Caches, insbes. bei archive.org
- ✓ Seit wann ist die Seite dort bekannt?
- ✓ Wie hat sie sich entwickelt?
- ✓ Vergleichen Sie die unterschiedlichen Versionen!



## Externe Kriterien (2)

### Versteckte Daten und Metadaten

#### HTTP-header

Der HTTP-Header zeigt die Kommunikation zwischen Webserver und Ihrem Browser, während die Seite geladen wird. Den können Sie mit einer Firefox-Erweiterung ansehen

[ → "Live HTTP Headers" und "Header Monitor" ]

- ✓ welcher Status-Code wird gesendet ("200 OK" ist normal)?
- ✓ was steht im Feld "Last-Modified"?
- ✓ wenn Sie eine Seite länger beobachten wollen, achten Sie auf "eTag" und "Content-Length"!



## Externe Kriterien (2)

### Versteckte Daten und Metadaten

#### Man kann immer noch weiter bohren...

- ✓ “ping” zeigt die IP-Adresse des Servers.
- ✓ “Trace” zeigt den Weg der Daten vom Server zu Ihrem Computer.
- ✓ ist die Seite speziell (dynamisch) für Sie erstellt?  
Webserver reagieren auf  
IP-Adressen / Session-IDs / Cookies / Referer / ...  
Surfen Sie ggfs. anonym!
- ✓ ist die Seite ein hoax? (Konsultieren Sie Hoax-Verzeichnisse!)
- ✓ ist Ihr DNS-Server (oder das des WLAN-Routers) gehacked?  
Sichern Sie Ihr System!





Vielen Dank für die Aufmerksamkeit!

Fragen dazu?

Entweder jetzt  
oder jederzeit an:

Albrecht Ude  
albrecht@ude.de  
www.ude.de

