



Hauptkriterien Suchindex und Benutzerinteraktion



Studienresultate: Evaluation der Suchfunktion
von Schweizer Unternehmens-Websites

Zürich, 30. November 2006
Jürg Stuker, namics

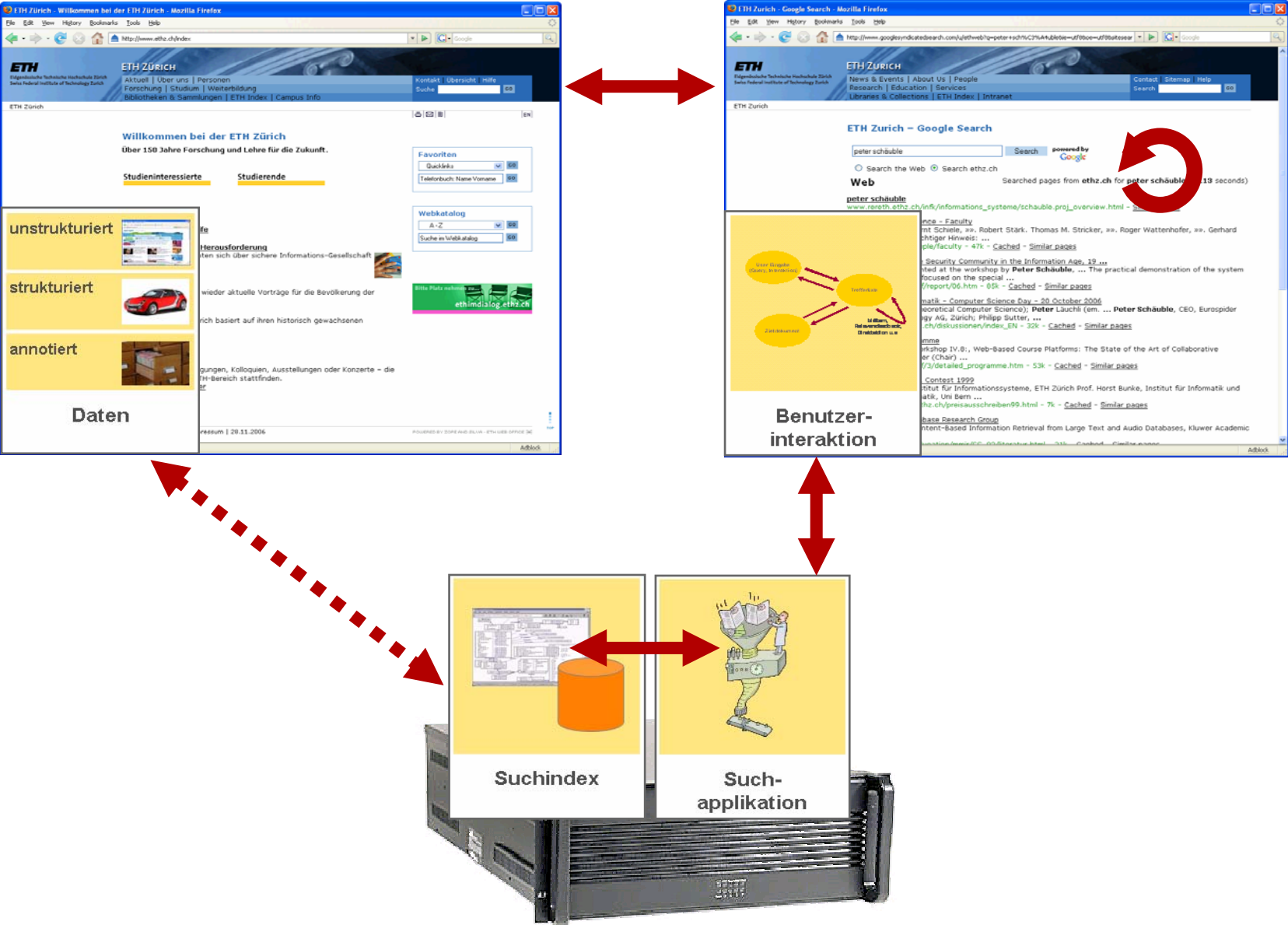
Agenda

- » Übersicht über zwei der Hauptkriterien
 - Suchindex
 - Benutzerinteraktion
- » Interpretation der Ergebnisse

Hauptkriterium Suchindex



Hauptkriterium Suchindex „in a nutshell“



Hauptkriterium Suchindex „in a nutshell“

- » Die Suchapplikation kann nur verarbeiten, was im Index vorhanden ist
 - Fehlt ein Dokument im Index, wird es nie auf der Suchtrefferliste erscheinen...
 - Ist ein Dokument fehlerhaft im Index, so wird die Suchanwendung kein korrektes Resultat liefern können...

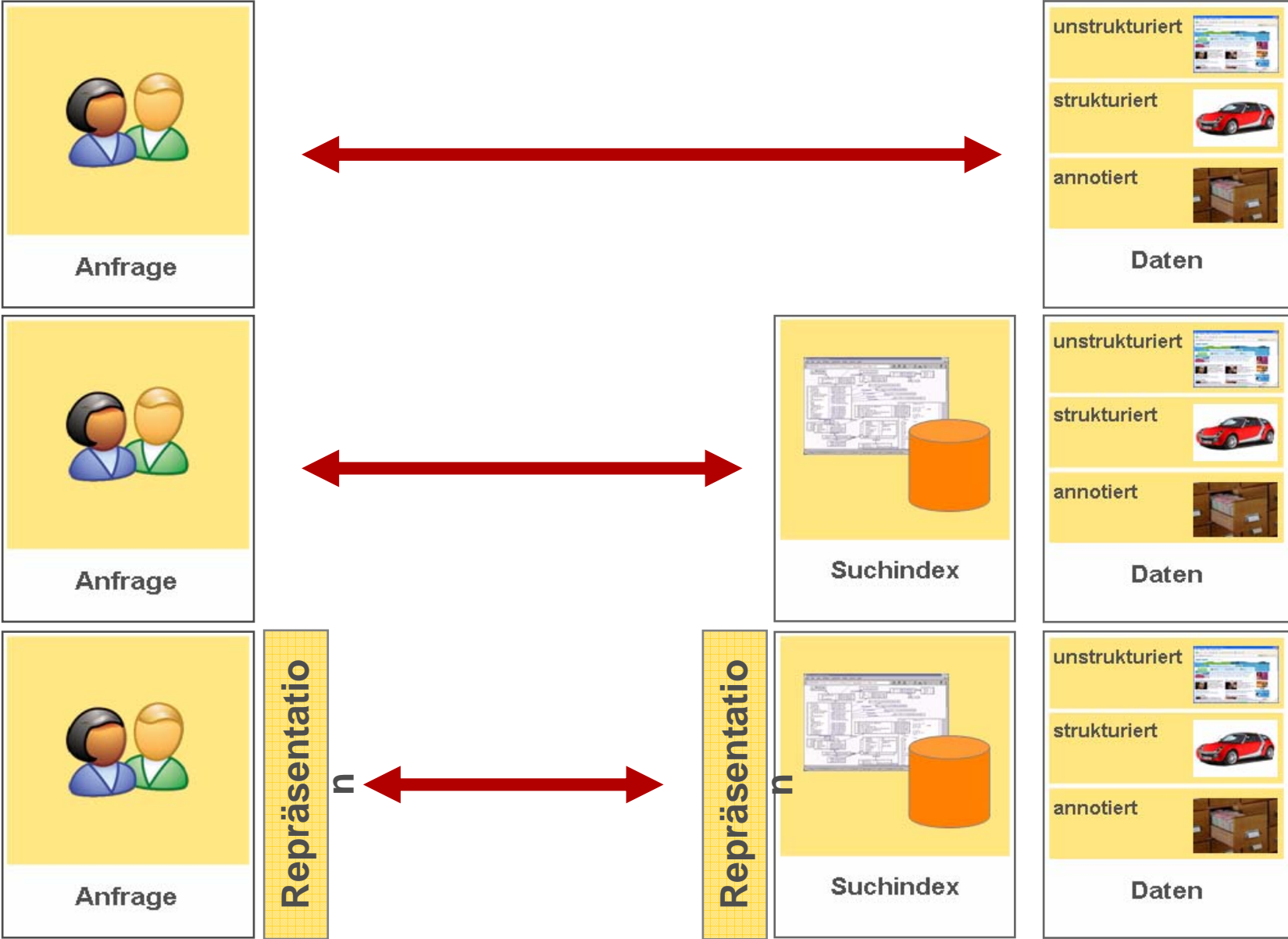
Subkriterien von Suchindex

- » Vollständigkeit
 - Sind alle möglichen Treffer der Suchtrefferliste im Index vorhanden?

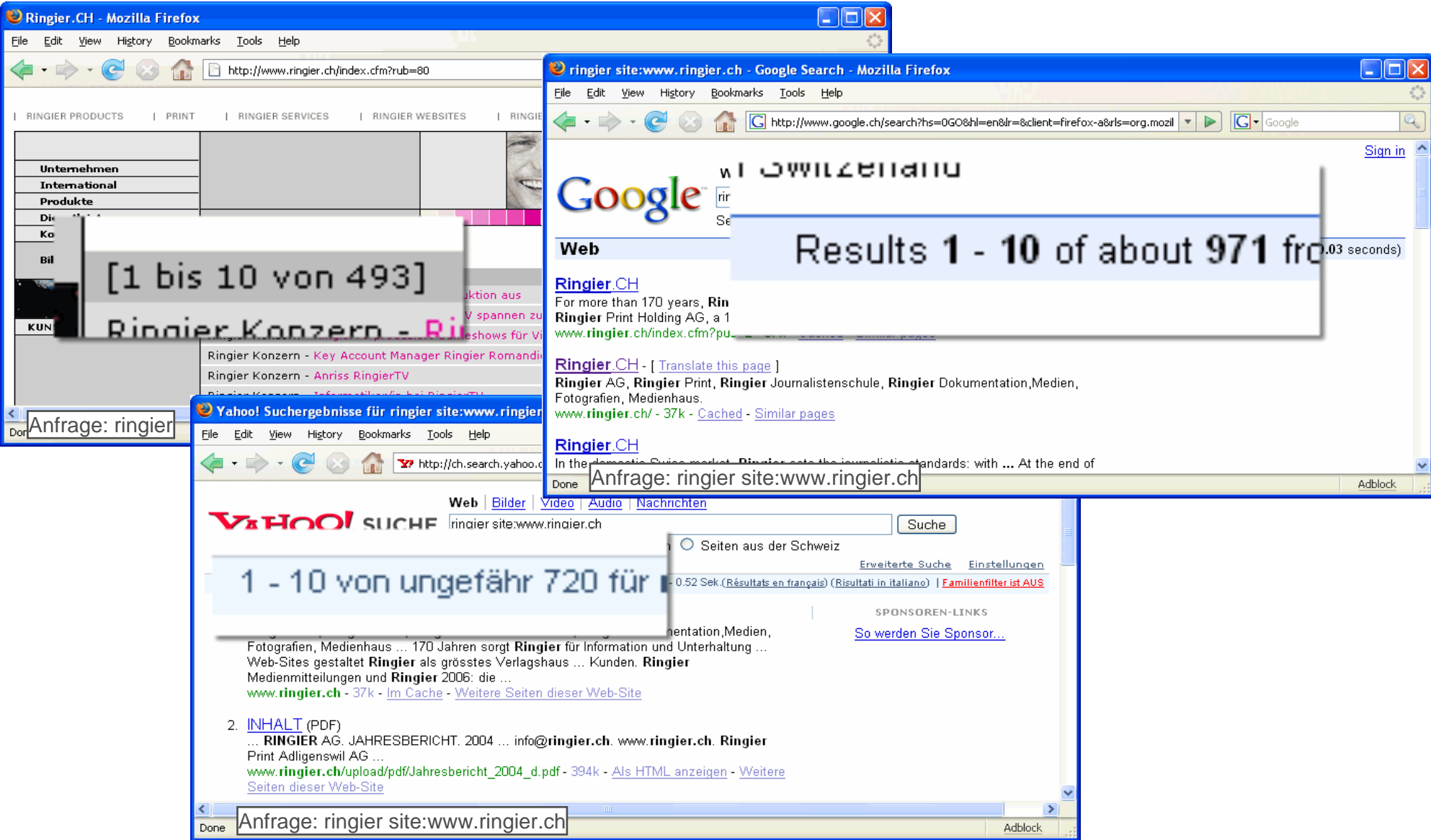
- » Aktualität
 - Existieren alle Seiten der Website als mögliche Treffer der Suchtrefferliste? [**create**]
 - Existieren alle möglichen Treffer der Suchtrefferliste auch in der (verlinkten) Quelle? [**delete**]
 - Stimmen alle möglichen Trefferzitate der Suchtrefferliste mit der aktuellen Quelle überein? [**update**]

- » Anfrage- / Dokumenterschliessung
 - Tokenisierung: Segmentierung von Text in linguistische Einheiten bspw. F/A-18 oder Hans-Peter O'Brian
 - Normalisierung: Behandlung von speziellen Zeichen und Schreibweisen wie bspw. von ä,ö,ü oder JÜRIG

Detail: Anfrage- / Dokumenterschliessung



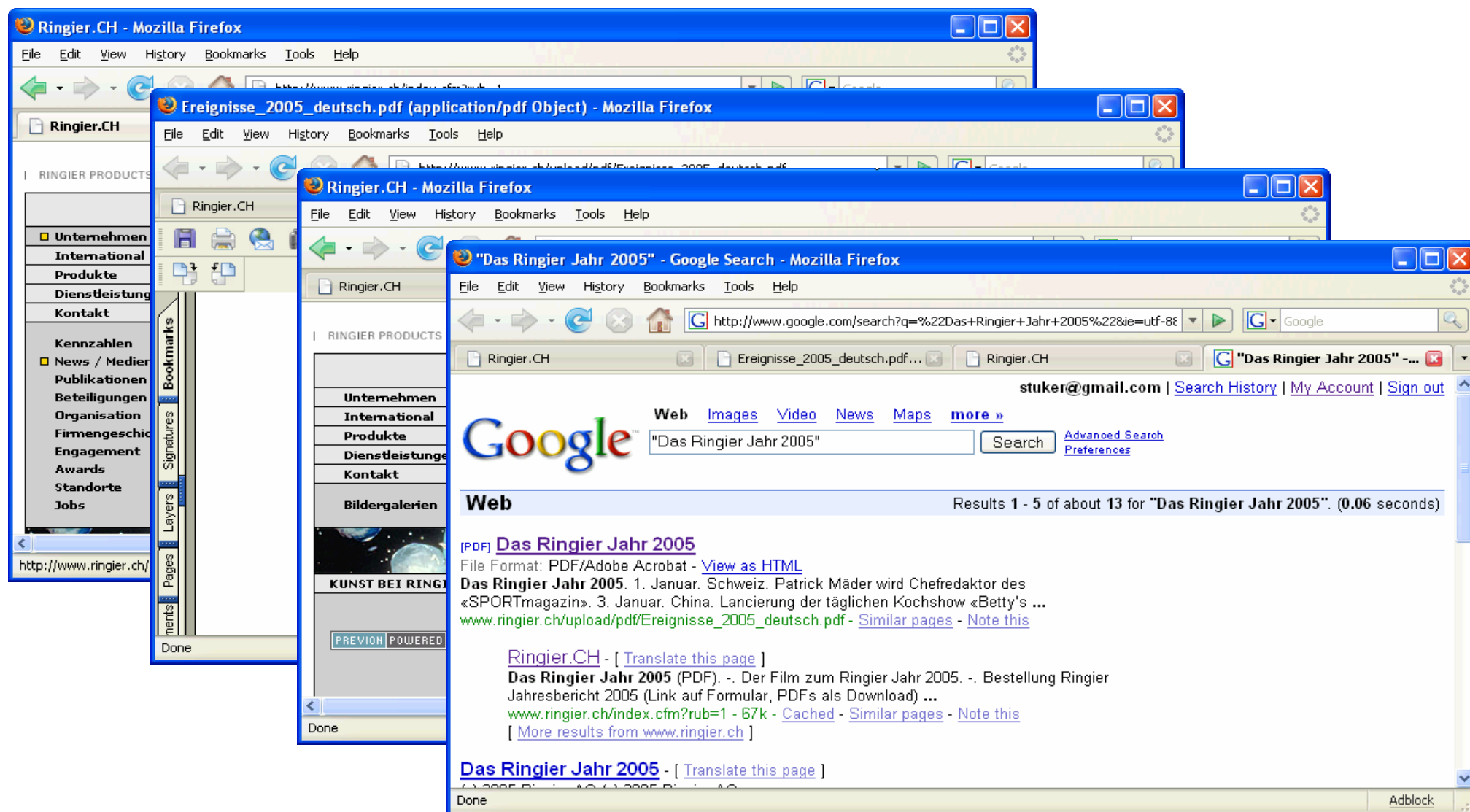
Beispiel Vollständigkeit (1 von 2)



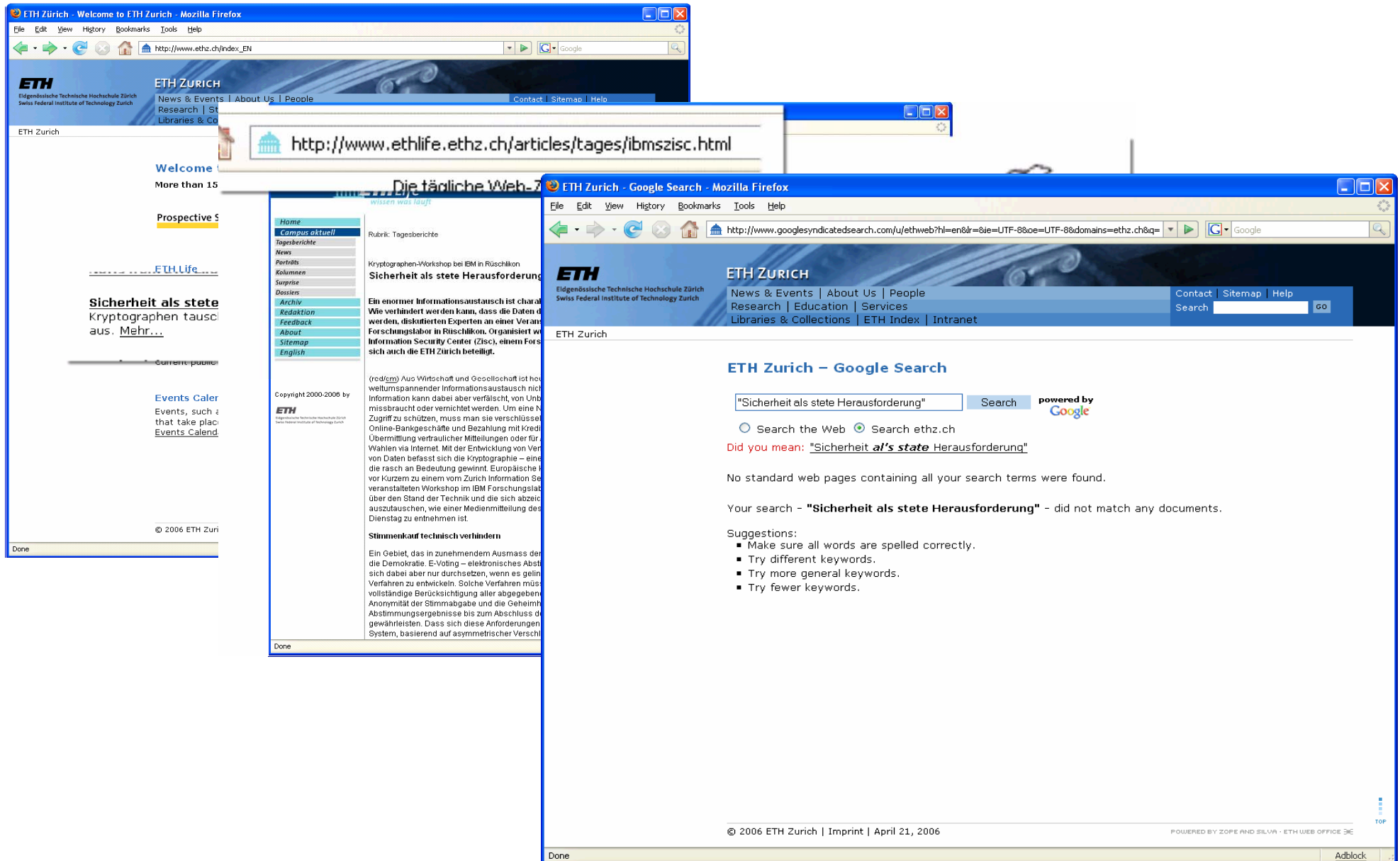
The image displays three overlapping browser windows illustrating search results for the query "ringier site:www.ringier.ch".

- Top-left window (Ringier.CH - Mozilla Firefox):** Shows the Ringier website's navigation menu with categories like "Unternehmen", "International", and "Produkte". A search bar contains the query "Anfrage: ringier" and shows results for "Ringier Konzern - Ri...".
- Top-right window (ringier site:www.ringier.ch - Google Search - Mozilla Firefox):** Shows Google search results for the query. The search bar contains "Anfrage: ringier site:www.ringier.ch". The results show "Results 1 - 10 of about 971 from...". The first result is "Ringier.CH" with a description: "For more than 170 years, Ringier Print Holding AG, a...".
- Bottom window (Yahoo! Suchergebnisse für ringier site:www.ringier.ch):** Shows Yahoo! search results for the query. The search bar contains "Anfrage: ringier site:www.ringier.ch". The results show "1 - 10 von ungefähr 720 für...". The first result is "Ringier.CH" with a description: "In the domestic Swiss market, Ringier sets the journalistic standards: with ... At the end of...".

Beispiel Vollständigkeit (2 von 2)

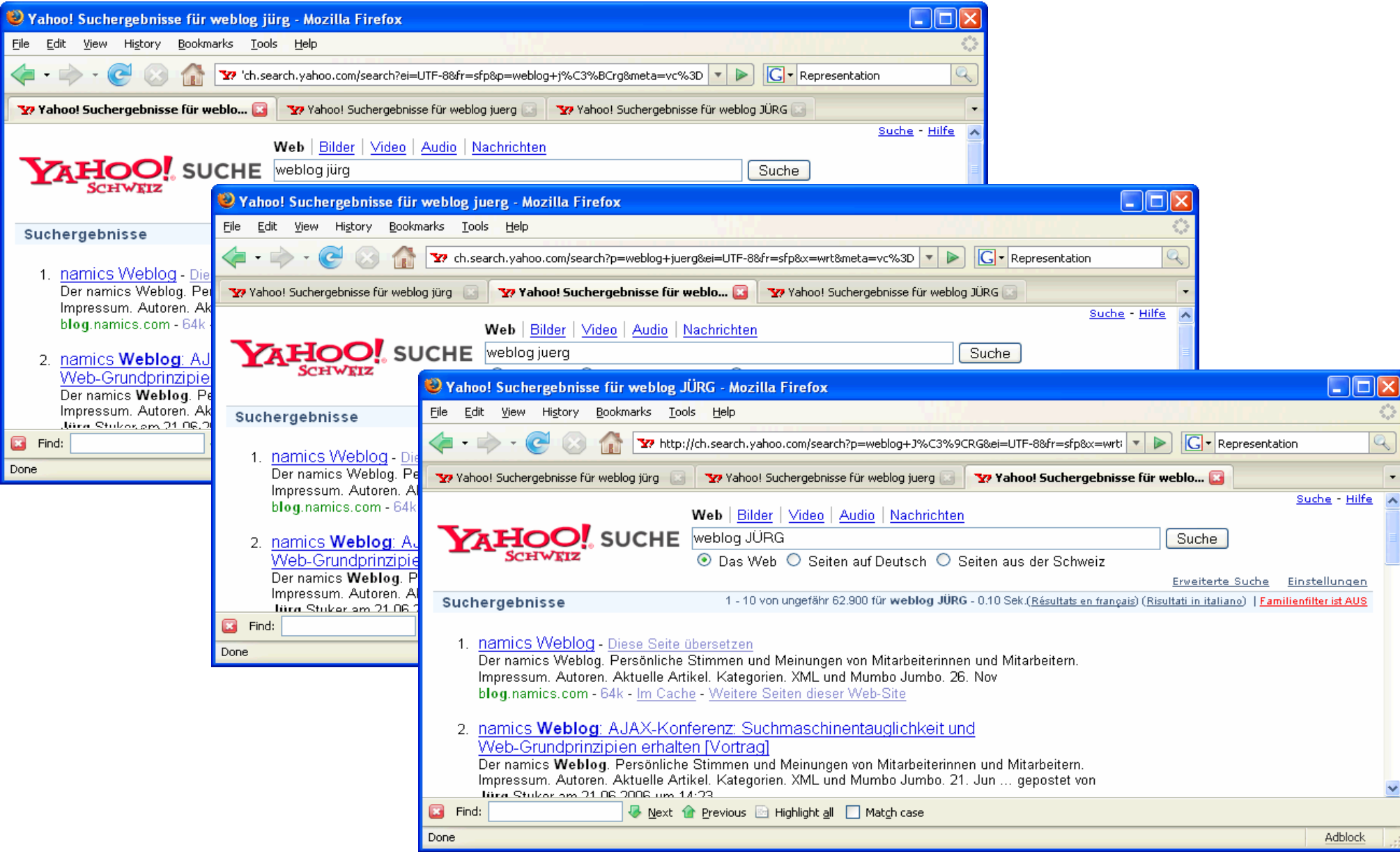


Beispiel Aktualität



The image shows a screenshot of a Mozilla Firefox browser window displaying the ETH Zurich website. The browser's address bar shows the URL http://www.ethz.ch/index_EN. The website header includes the ETH Zurich logo and navigation links such as 'News & Events', 'About Us', and 'People'. A sidebar on the left contains a menu with items like 'Home', 'Campus aktuell', 'Tagesberichte', 'News', 'Porträts', 'Kolumnen', 'Surprise', 'Dossiers', 'Archiv', 'Redaktion', 'Feedback', 'About', 'Sitemap', and 'English'. The main content area features a section titled 'Sicherheit als stete Herausforderung' with a sub-heading 'Kryptographen-Workshop bei IBM in Rüschlikon'. The text discusses the challenges of information security in the digital age, mentioning the IBM Research Security Center (Zisc) and the importance of secure communication. A search engine window is overlaid on the right, showing a search for 'Sicherheit als stete Herausforderung' on the Google search engine. The search results indicate that no standard web pages containing all the search terms were found, and it suggests checking for typos or using more general keywords.

Bps. Anfrage- / Dokumenterschliessung (1 von 2)



The image displays three overlapping screenshots of a Mozilla Firefox browser window showing Yahoo! search results. The browser's address bar and search bar are visible in each screenshot, along with the search results list.

Top Screenshot: Yahoo! Suche für weblog jürg

- Search query: `ch.search.yahoo.com/search?ei=UTF-8&fr=sfp&p=weblog+j%C3%BCrg&meta=vc%3D`
- Search bar: `weblog jürg`
- Search results:
 - [namics Weblog](#) - Die Der namics Weblog. Pe Impressum. Autoren. Ak [blog.namics.com](#) - 64k
 - [namics Weblog: AJ Web-Grundprinzipie](#) Der namics **Weblog**. Pe Impressum. Autoren. Ak [lira Stuker am 21.06.2006](#)

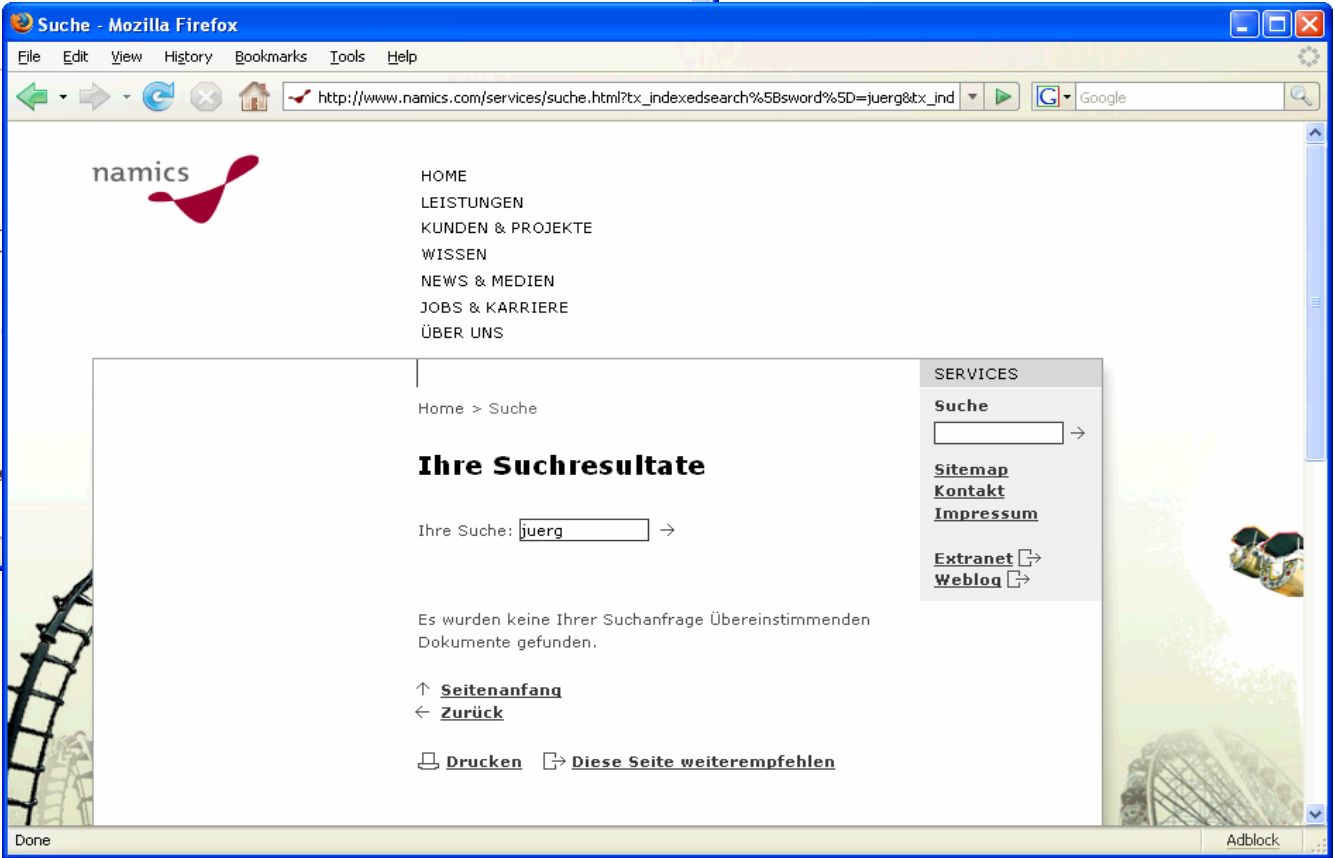
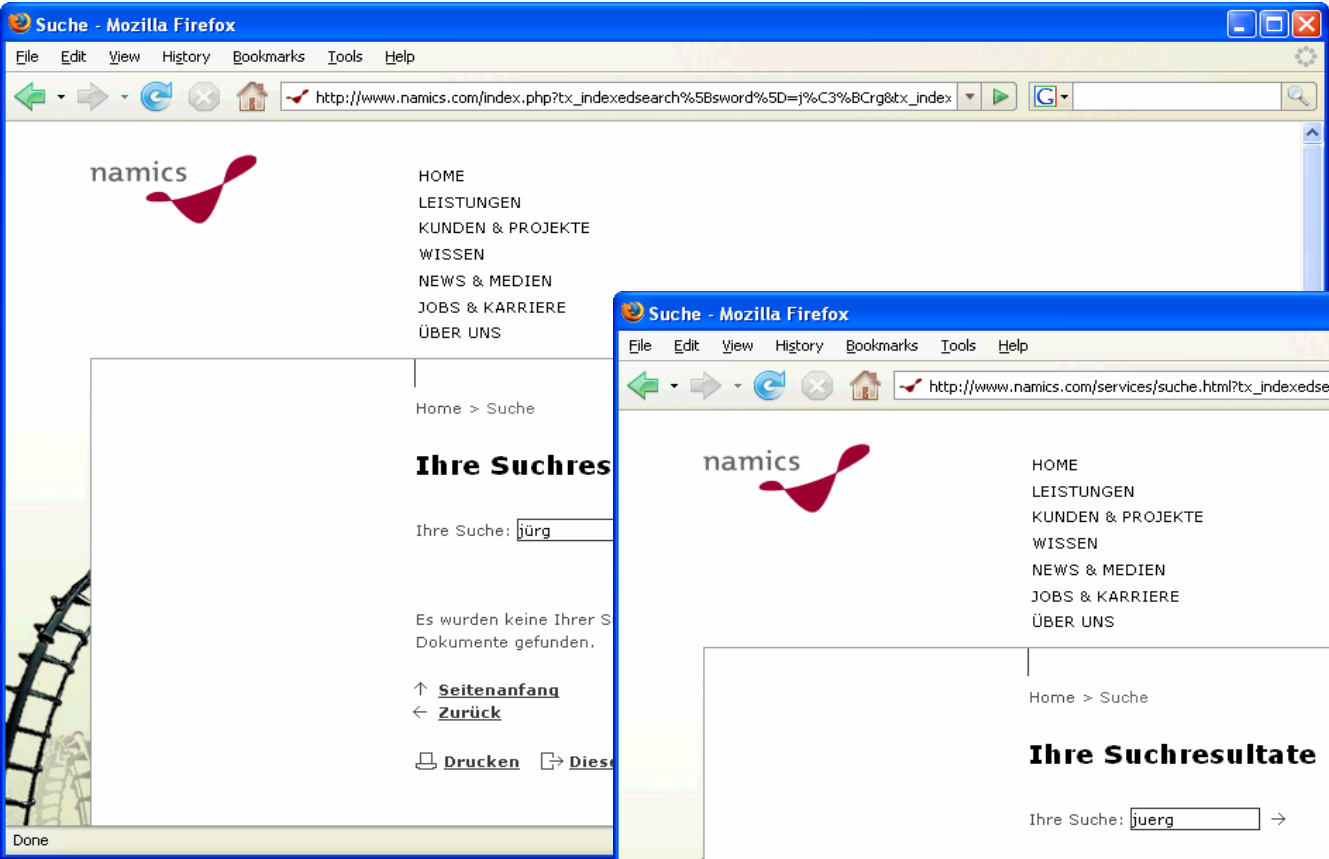
Middle Screenshot: Yahoo! Suche für weblog juerg

- Search query: `ch.search.yahoo.com/search?p=weblog+juerg&ei=UTF-8&fr=sfp&x=wrt&meta=vc%3D`
- Search bar: `weblog juerg`
- Search results:
 - [namics Weblog](#) - Die Der namics Weblog. Pe Impressum. Autoren. Ak [blog.namics.com](#) - 64k
 - [namics Weblog: AJ Web-Grundprinzipie](#) Der namics **Weblog**. Pe Impressum. Autoren. Ak [lira Stuker am 21.06.2006](#)

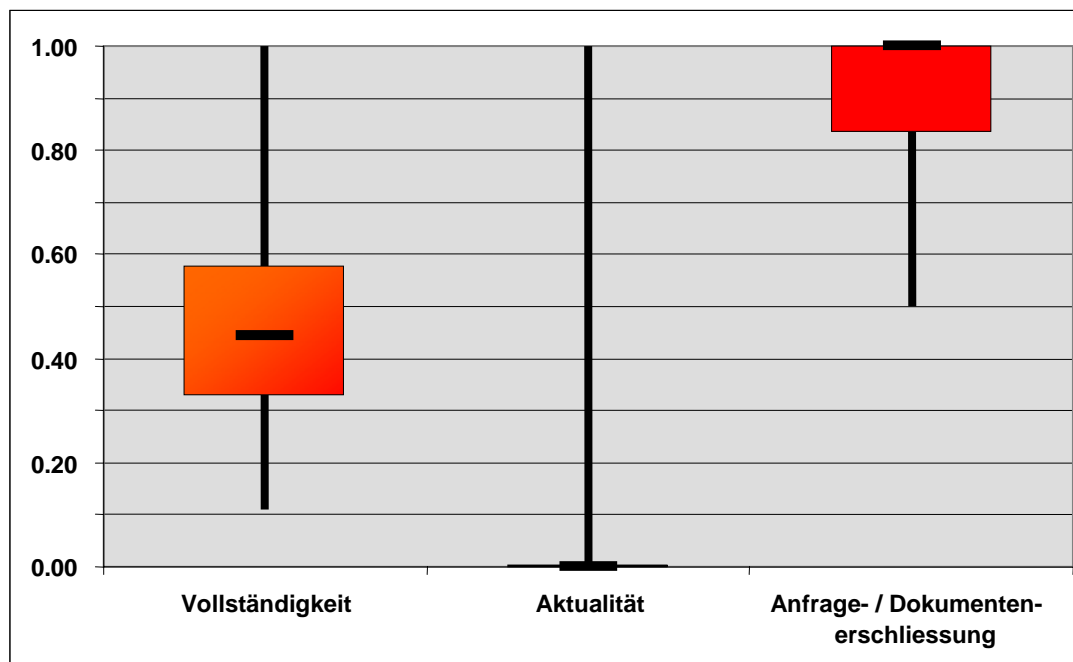
Bottom Screenshot: Yahoo! Suche für weblog JÜRG

- Search query: `http://ch.search.yahoo.com/search?p=weblog+J%C3%9CRG&ei=UTF-8&fr=sfp&x=wrt`
- Search bar: `weblog JÜRG`
- Search results:
 - [namics Weblog](#) - [Diese Seite übersetzen](#) Der namics Weblog. Persönliche Stimmen und Meinungen von Mitarbeiterinnen und Mitarbeitern. Impressum. Autoren. Aktuelle Artikel. Kategorien. XML und Mumbo Jumbo. 26. Nov [blog.namics.com](#) - 64k - [Im Cache](#) - [Weitere Seiten dieser Web-Site](#)
 - [namics Weblog: AJAX-Konferenz: Suchmaschinentauglichkeit und Web-Grundprinzipien erhalten \[Vortrag\]](#) Der namics **Weblog**. Persönliche Stimmen und Meinungen von Mitarbeiterinnen und Mitarbeitern. Impressum. Autoren. Aktuelle Artikel. Kategorien. XML und Mumbo Jumbo. 21. Jun ... gepostet von [lira Stuker am 21.06.2006 um 14:23](#)

Bps. Anfrage- / Dokumenterschliessung (2 von 2)



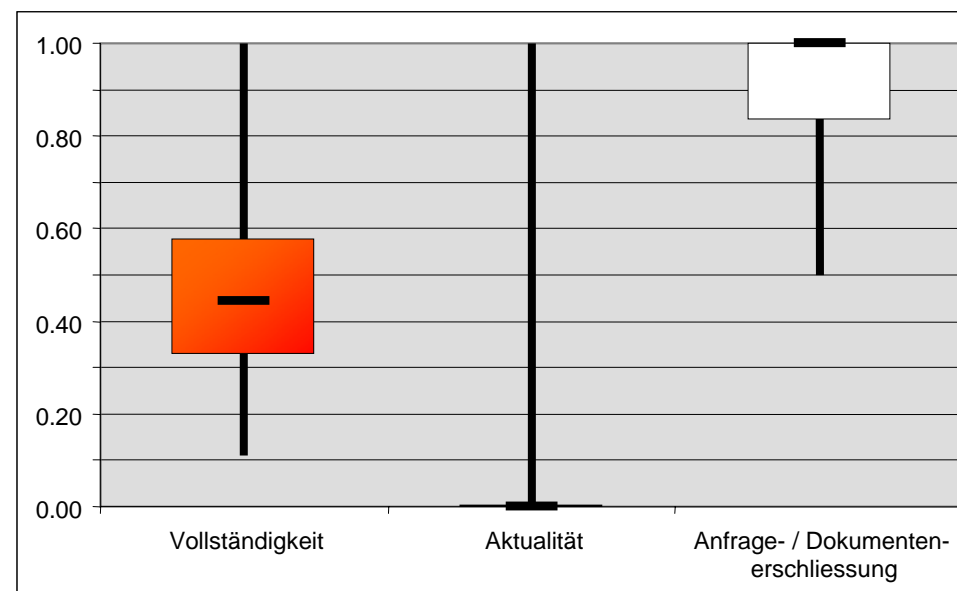
Resultatüberblick Hauptkriterium Suchindex



	Vollständigkeit	Aktualität	Anfrage-/Dokument-erschliessung
Maximum	1.00	1.00	1.00
Median	0.44	0.00	1.00
Minimum	0.11	0.00	0.50
IQ-Abstand	0.25	0.00	0.17

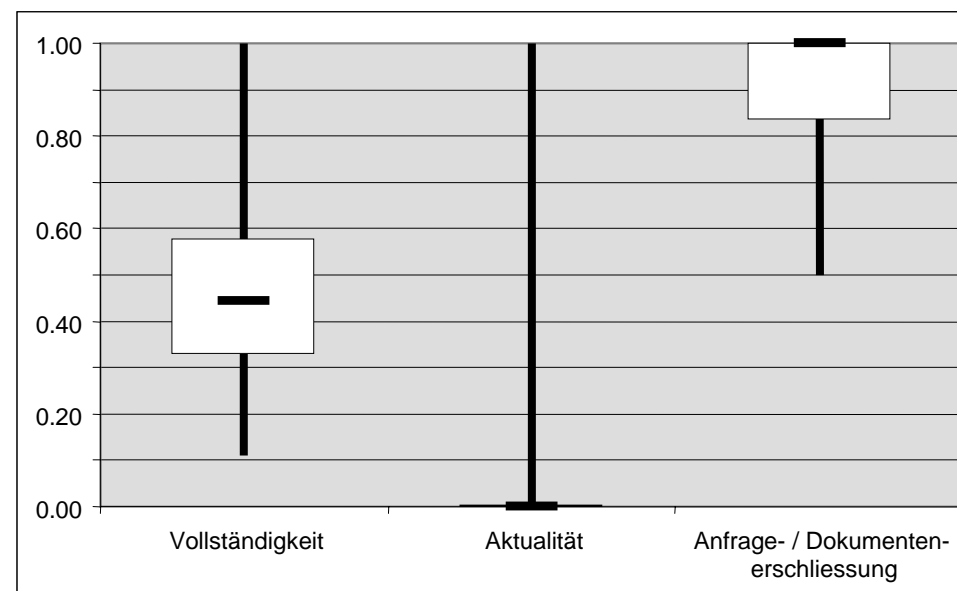
Interpretation Vollständigkeit

- » Median ist tief
 - Dokumentenzahl ist bekannt
 - Dokumente sind öffentlich
 - Alle Elemente sind im Wirkungsbereich des Betreibers
- » Maximum zeigt: Aufgabe ist lösbar (eine Site, dafür nicht aktuell ;-)
- » Regelmässige Probleme
 - PDF-Dokumente: Erschliessung und Textextraktion
 - Anwendungen sind nicht erschlossen (HTML Formulare)



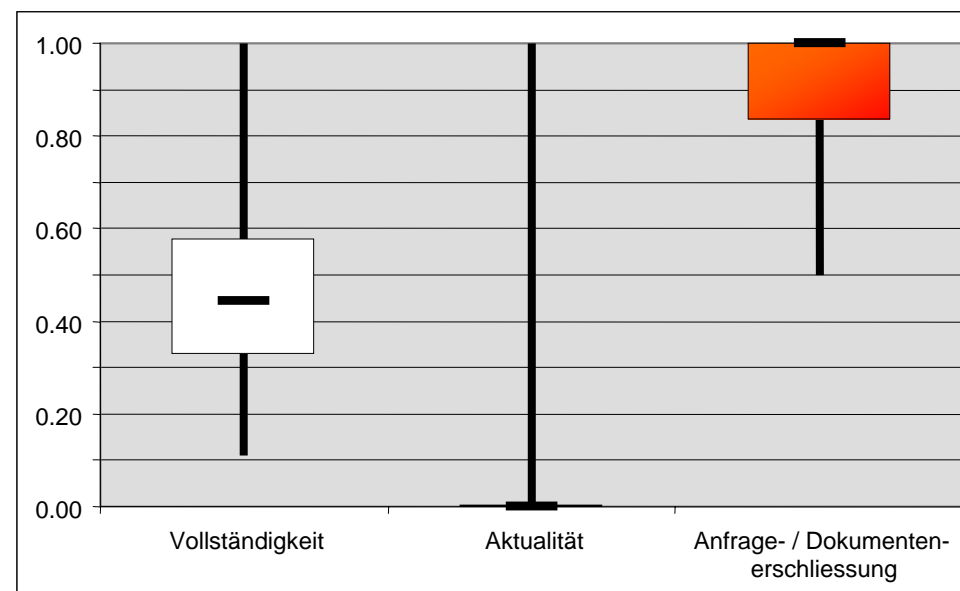
Interpretation Aktualität

- » Median ist Null
 - Jeder Fehlertyp war häufig
- » Maximum zeigt: Aufgabe ist lösbar (drei Sites)
- » Regelmässige Probleme
 - Homepage-Inhalte sind nicht erschlossen (News)
 - Links zu nicht existierende Seiten in der Trefferliste (http 404)



Interpretation Anfrage- / Dokumenterschliessung

- » Median ist Eins
- » Streuung erfreulich gering
- » 29 Sites erreichten das Maximum
- » Die Betreiber meistern die Mehrsprachigkeit in der Schweiz
- » Normalisierung (44 Sites) wurde häufiger erfüllt als Tokenisierung (34 Sites)



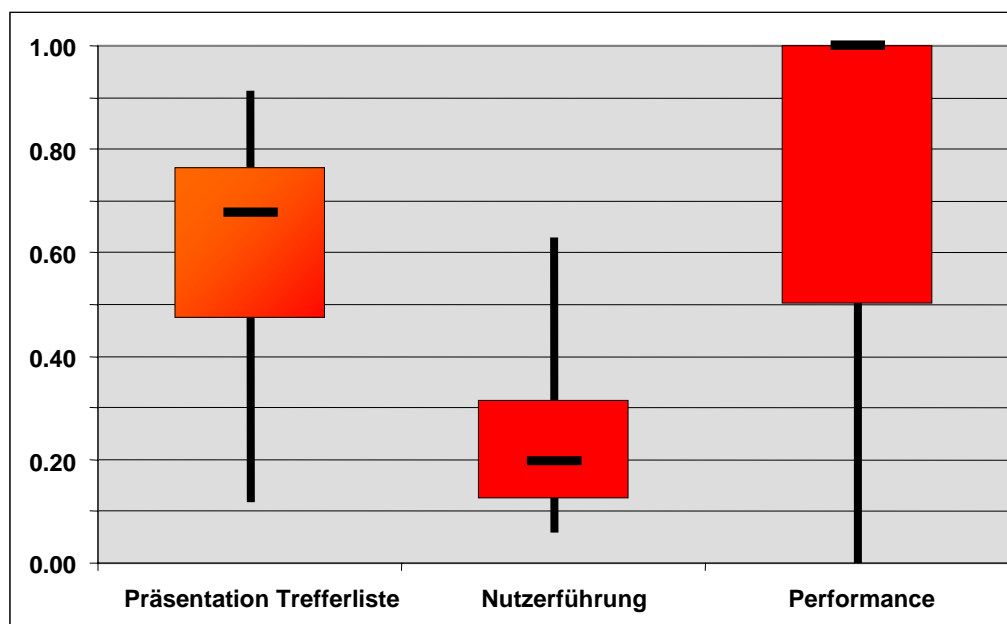
Hauptkriterium Benutzerinteraktion



Hauptkriterium Benutzerinteraktion „in a nutshell“

- » Das, was der Nutzer von der Suchfunktion sieht und bedient...
- » Präsentation der Trefferliste
 - Qualität der Trefferliste ohne Interaktion inkl. Trefferzeit und Bestandteile eines Treffers
 - Verständlichkeit der visuellen Codierung
- » Nutzerführung
 - Verfügbarkeit und Qualität Funktionen / Interaktion
 - Aktive Fehlervermeidung und/oder Nutzerführung
 - Subkollektionen, Behandlung von Null Treffern und Tippfehlern...
- » Performance

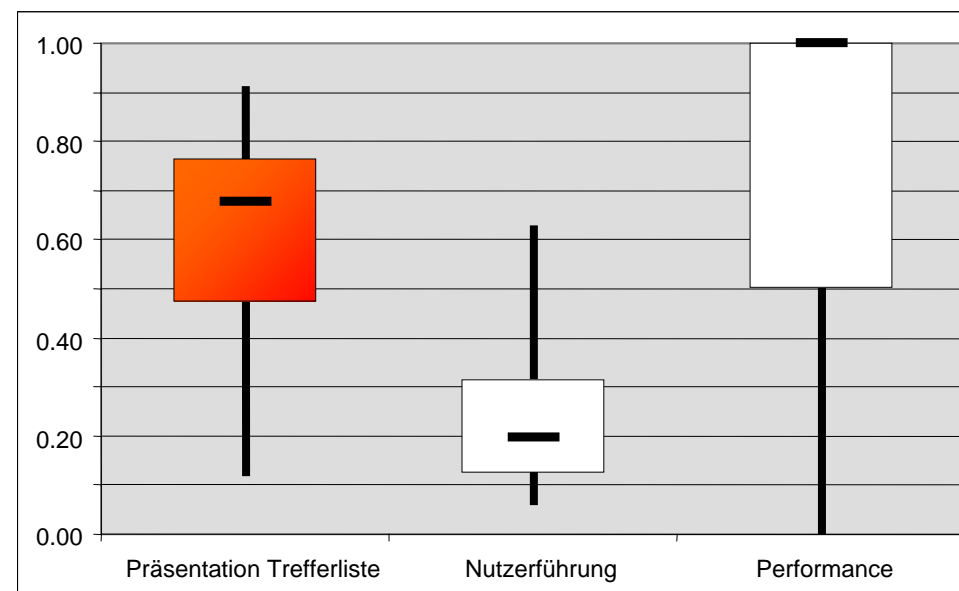
Resultatüberblick Hauptkriterium Benutzerinteraktion



	Präsentation Trefferliste	Nutzerführung	Performance
Maximum	0.91	0.63	1.00
Median	0.68	0.20	1.00
Minimum	0.12	0.06	0.00
IQ-Abstand	0.29	0.19	0.50

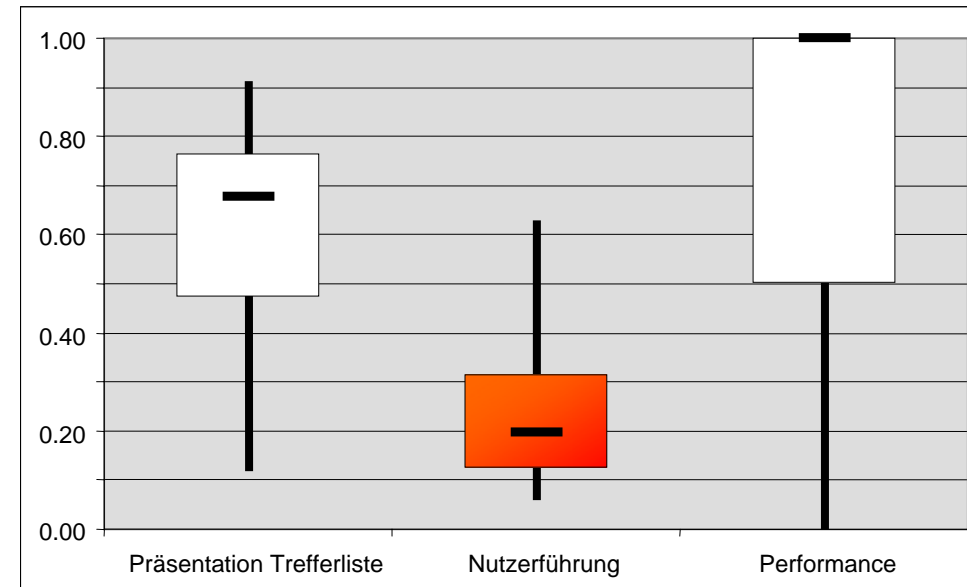
Interpretation Präsentation der Trefferliste

- » Median ist mit 0.68 ziemlich gut
- » Varianz ist gross
- » Erstaunlich da taugliche Vorbilder vorhanden sind
- » Häufige Probleme
 - Trefferzitate (Vorhandensein und/oder Qualität)
 - PDF-Dokumente: Trefferzitate, Cache-Versionen und Highlighting (nur 8 Sites)
 - Metadaten: Dokumentgrösse (22 Sites), Dokumentdatum (24 Sites) und Dokumenttyp (26 Sites)



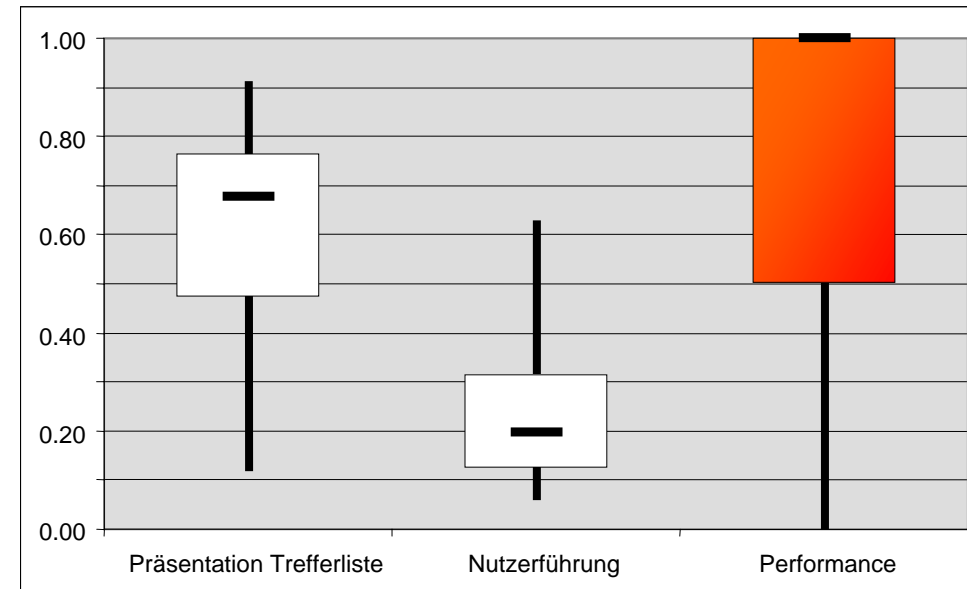
Interpretation Nutzerführung

- » Median mit 0.2 sehr tief und nahe beim Minimum
- » Auch erster (0.12) und dritter Quartilswert (0.32) sind tief
- » Ein Teil der Funktionen ist nicht ganz einfach zu implementieren
- » Aber viele der Qualitätsaspekte wären einfach zu lösen: Anstelle Nutzerfokus eher technischer Fokus bei der Umsetzung?
- » Häufige Probleme: Fast alles...
 - Tippfehler, weitere Suchbegriffe vorschlagen, Subkollektionen, Suche nach ähnlichen Dokumenten u.S.W.

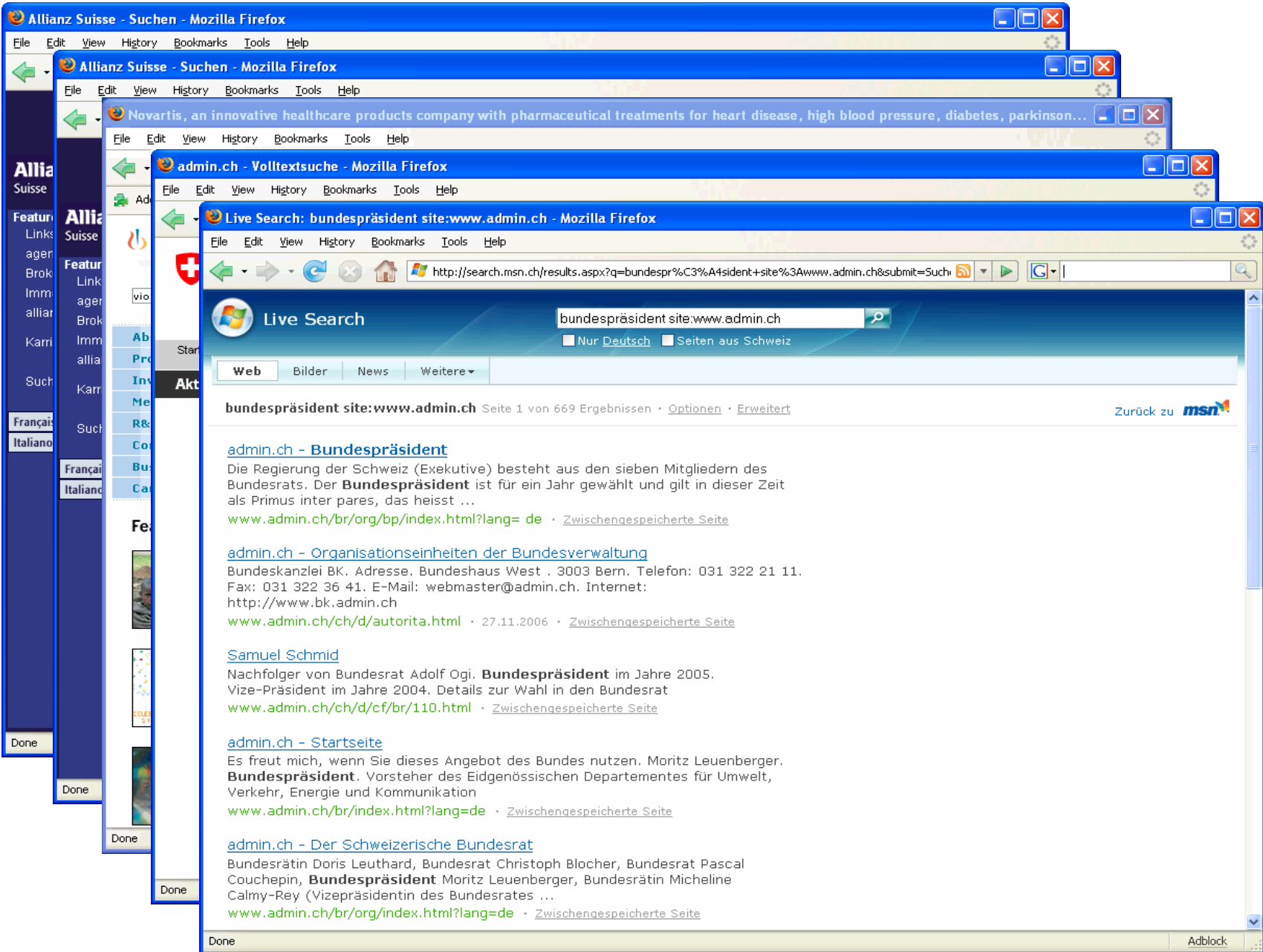


Interpretation Performance

- » Gutes Ergebnis
- » Augenfällig, dass eine gute Performance kaum beachtet wird, sehr wohl aber eine schlechte
- » Suchapplikationen welche schlecht abschnitten, zeigten einen Zusammenhang der Geschwindigkeit mit der Länge der Anfrage



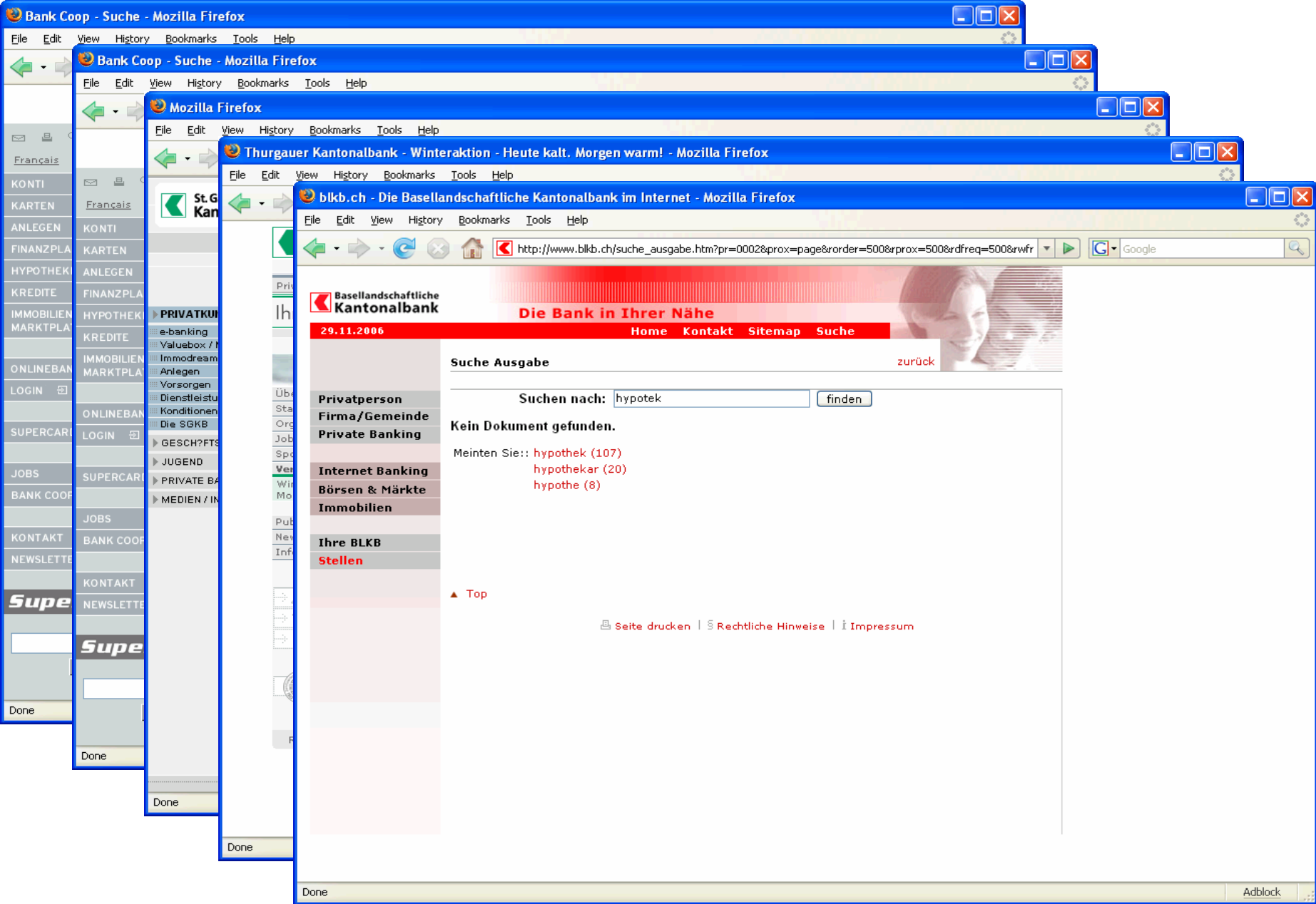
Beispiel Präsentation



The screenshot shows a Mozilla Firefox browser window with the following content:

- Address bar: `http://search.msn.ch/results.aspx?q=bundespr%C3%A4sident+site%3Awww.admin.ch&submit=Suche`
- Search engine: Live Search
- Search query: `bundespräsident site:www.admin.ch`
- Filters: Nur Deutsch, Seiten aus Schweiz
- Search results: `bundespräsident site:www.admin.ch` Seite 1 von 669 Ergebnissen
- Results list:
 - [admin.ch - Bundespräsident](#)
Die Regierung der Schweiz (Exekutive) besteht aus den sieben Mitgliedern des Bundesrats. Der **Bundespräsident** ist für ein Jahr gewählt und gilt in dieser Zeit als *Primus inter pares*, das heisst ...
[www.admin.ch/br/org/bp/index.html?lang=de](#) · [Zwischengespeicherte Seite](#)
 - [admin.ch - Organisationseinheiten der Bundesverwaltung](#)
Bundeskanzlei BK. Adresse. Bundeshaus West . 3003 Bern. Telefon: 031 322 21 11. Fax: 031 322 36 41. E-Mail: webmaster@admin.ch. Internet: <http://www.bk.admin.ch>
[www.admin.ch/ch/d/autorita.html](#) · 27.11.2006 · [Zwischengespeicherte Seite](#)
 - [Samuel Schmid](#)
Nachfolger von Bundesrat Adolf Ogi. **Bundespräsident** im Jahre 2005. Vize-Präsident im Jahre 2004. Details zur Wahl in den Bundesrat
[www.admin.ch/ch/d/cf/br/110.html](#) · [Zwischengespeicherte Seite](#)
 - [admin.ch - Startseite](#)
Es freut mich, wenn Sie dieses Angebot des Bundes nutzen. Moritz Leuenberger. **Bundespräsident**. Vorsteher des Eidgenössischen Departementes für Umwelt, Verkehr, Energie und Kommunikation
[www.admin.ch/br/index.html?lang=de](#) · [Zwischengespeicherte Seite](#)
 - [admin.ch - Der Schweizerische Bundesrat](#)
Bundesrätin Doris Leuthard, Bundesrat Christoph Blocher, Bundesrat Pascal Couchepin, **Bundespräsident** Moritz Leuenberger, Bundesrätin Micheline Calmy-Rey (Vizepräsidentin des Bundesrates ...
[www.admin.ch/br/org/index.html?lang=de](#) · [Zwischengespeicherte Seite](#)

Beispiel Nutzerführung



Danke für Ihre Aufmerksamkeit.



juerg.stuker@namics.com
http://blog.namics.com/information_retrieval/