

## b-14 Enterprise Search: Google und Microsoft

Freitag, 23. Mai 2008

René Knöfel, Technical Solution Professional, Microsoft GmbH

Joachim Glaubrecht, Product Specialist, Google

Marco Billeter, Director Central Europe, Fast

Jürg Stuker, CEO & Partner, namics ag

Bern, Frankfurt, Hamburg, München, St. Gallen, Zug, Zürich

1

### Zielsetzungen

- » Einblick in die branchenführenden Lösungen von Microsoft (und Fast) sowie von Google ermöglichen
- » Ideen und Tipps mitgeben um Ihre Suchanwendung besser zu machen
- » Kontakt zu Experten ermöglichen, die sich täglich mit der Unternehmenssuche befassen
- » Abgrenzung (öffentliche) Websuche und Unternehmenssuche
- » Soviel wie sinnvoll am "echten Produkt" zeigen

2

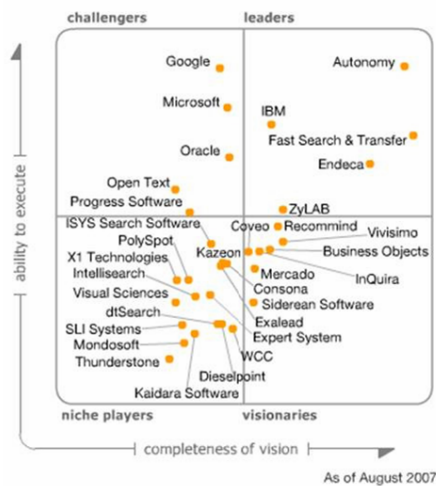
www.namics.com

## Herzlichen Dank an...

- » René Knöfel, Microsoft
- » Joachim Glaubrecht, Google
- » Marco Billeter, Fast



## Konsolidierung der Anbieter schreitet voran



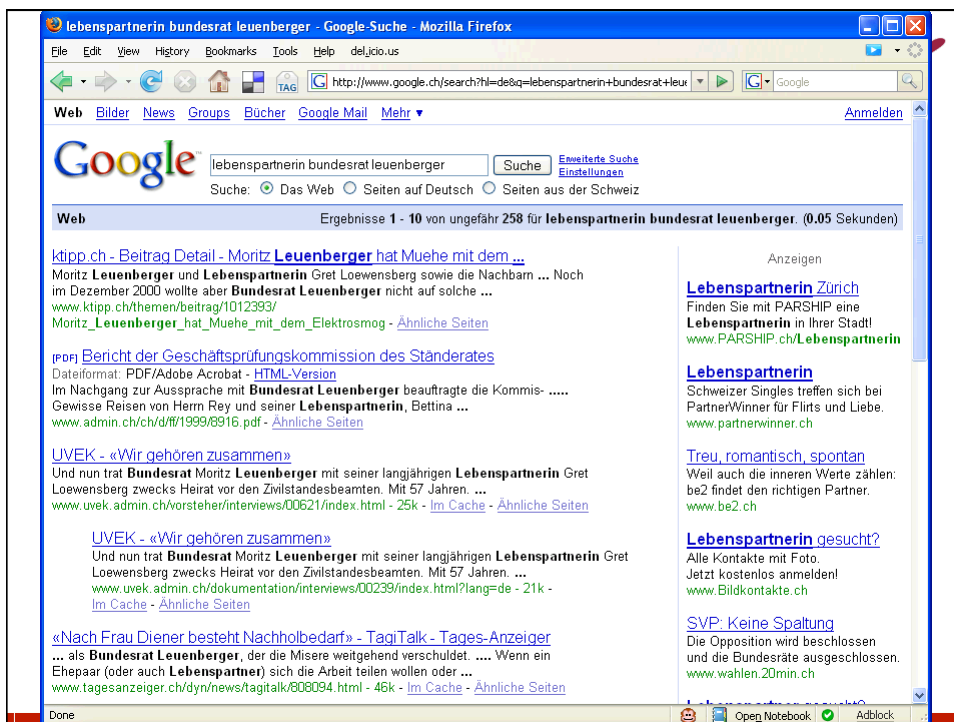
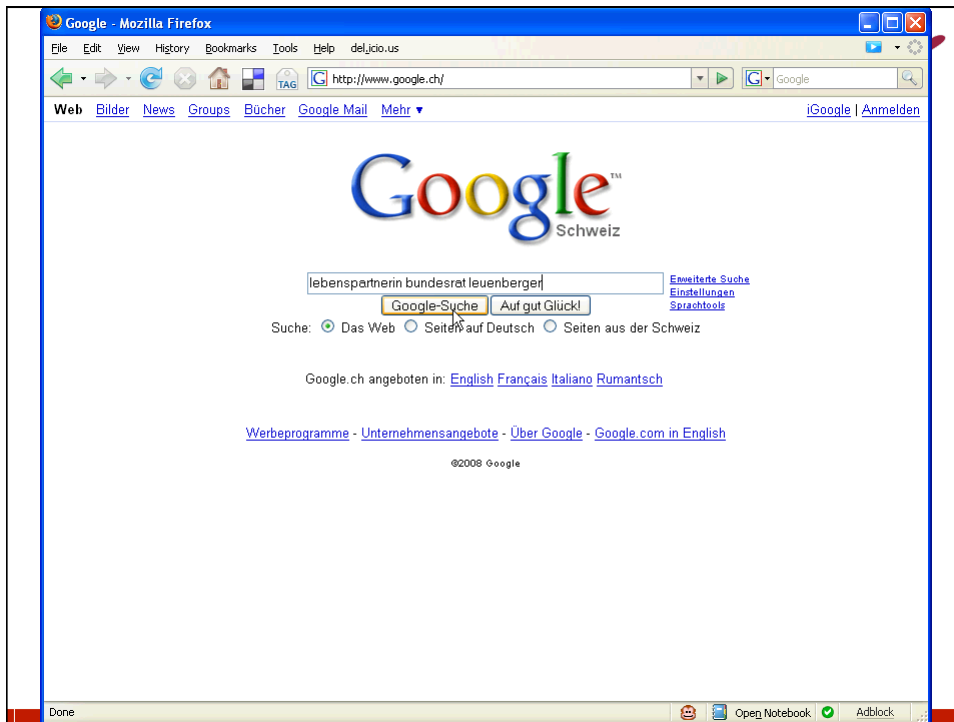
Quelle: Gartner Research

## Agenda

» Herausforderungen der unternehmensweiten Suche	namics	10'
» Vorstellung des Produktportfolios	Google Microsoft	je 10'
» Eine Trefferliste mit unterschiedlichen Quellen und Berechtigungen	Microsoft Google	je 10'
» Qualität der Rangliste	Google Microsoft	je 10'
» Tipps bei der Einführung einer unternehmensweiten Suche	namics	10'
» Fragen & Antworten	alle	10'

## Suchen sieht so einfach aus...

- » Sie haben eine Frage...



## Suchen sieht so einfach aus...

- » ...und Sie fragen sich vielleicht, weshalb das bei Ihnen im Intranet nicht genau so funktioniert?

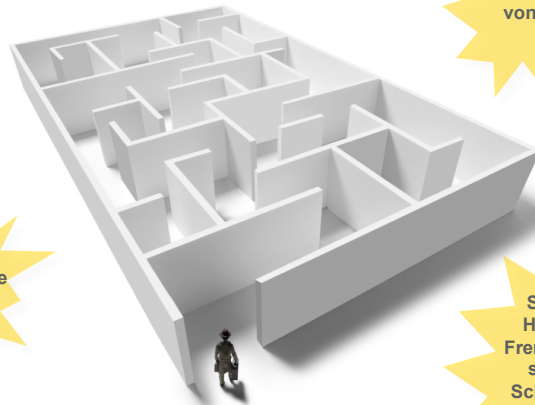
## Herausforderungen der unternehmensweiten Suche

100% namics 

## Herausforderungen bei der Suche

**I.**  
Unterschiedliche  
Arten von  
Suchaktivitäten

**II.**  
Unterschiedliche  
suchende  
Personen



**III.**  
Unterschiedliche  
Arten  
von Daten

**IV.**  
Synonyme,  
Homonyme,  
Fremdsprachen,  
schwierige  
Schreibweisen

## Kategorien von Suchaktivitäten

### "Lookup" (nachschnellen)

- Faktensuche
- Known Item Search
- Navigation
- Transaktion
- ...

### "Learn" (erforschen)

- Wissenserarbeitung
- Social Browsing
- Vergleiche
- Interpretation
- ...

### "Investigate" (untersuchen)

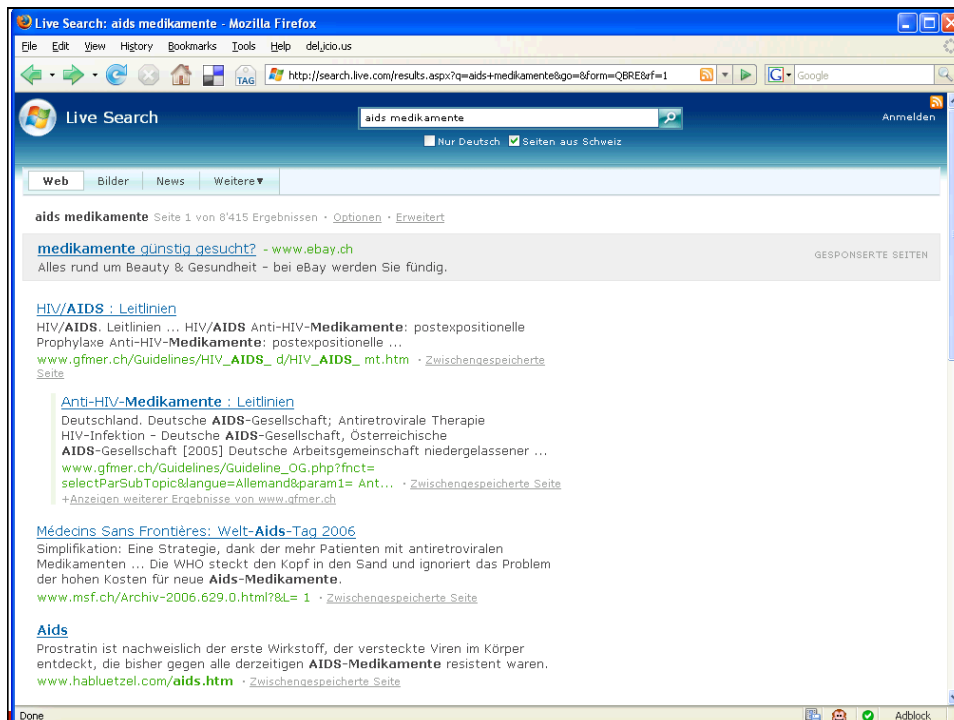
- Aggregation
- Analyse über Zeit
- Synthese
- Forecasting
- ...

» Die von der Suche erwartete Unterstützung ist je nach Aufgabe sehr unterschiedlich...

Quelle: Gary Marchionini, From Finding to Understanding, CACM April 2006/Vol. 49, No. 4

## Suchende Personen

- » Menschen sind unterschiedlich
  - so auch bei der Nutzung der Suche
- » Suchexperten versus “Suchlaien”
- » Retrievalprobleme
  - Ich suche, da ich die Antwort nicht kenne
  - Kenntnis der Antwort würde mir aber helfen (z. B. Zieldokument ist auf französisch)
  - Interpretation der gefundenen Information variiert von Nutzer zu Nutzer resp. während des Suchprozesses wird Domänenwissen erworben



Live Search: aids medikamente - Mozilla Firefox

File Edit View History Bookmarks Tools Help deljcio.us

http://search.live.com/results.aspx?q=aids+medikamente&go=&form=QBRE&f=1

Live Search  Anmelden

Nur Deutsch  Seiten aus Schweiz

Web Bilder News Weitere ▾

aids medikamente Seite 1 von 8'415 Ergebnissen · Optionen · Erweitert

[medikamente günstig gesucht?](#) - [www.ebay.ch](#) GESPONSERTE SEITEN

Alles rund um Beauty & Gesundheit - bei eBay werden Sie fündig.

[HIV/AIDS : Leitlinien](#)

HIV/AIDS Leitlinien ... HIV/AIDS Anti-HIV-Medikamente: postexpositionelle Prophylaxe Anti-HIV-Medikamente: postexpositionelle ...

[www.gfmer.ch/Guidelines/HIV\\_AIDS\\_d/HIV\\_AIDS\\_mt.htm](#) · [Zwischengespeicherte Seite](#)

[Anti-HIV-Medikamente : Leitlinien](#)

Deutschland, Deutsche AIDS-Gesellschaft; Antiretrovirale Therapie HIV-Infektion - Deutsche AIDS-Gesellschaft, Österreichische AIDS-Gesellschaft [2005] Deutsche Arbeitsgemeinschaft niedergelassener ...

[www.gfmer.ch/Guidelines/Guideline\\_OG.php?fnct=selectParSubTopic&langue=Allemand&param1= Ant...](#) · [Zwischengespeicherte Seite](#)

+Anzeigen weiterer Ergebnisse von [www.gfmer.ch](#)

[Médecins Sans Frontières: Welt-Aids-Tag 2006](#)


Simplifikation: Eine Strategie, dank der mehr Patienten mit antiretroviralen Medikamenten ... Die WHO steckt den Kopf in den Sand und ignoriert das Problem der hohen Kosten für neue Aids-Medikamente.

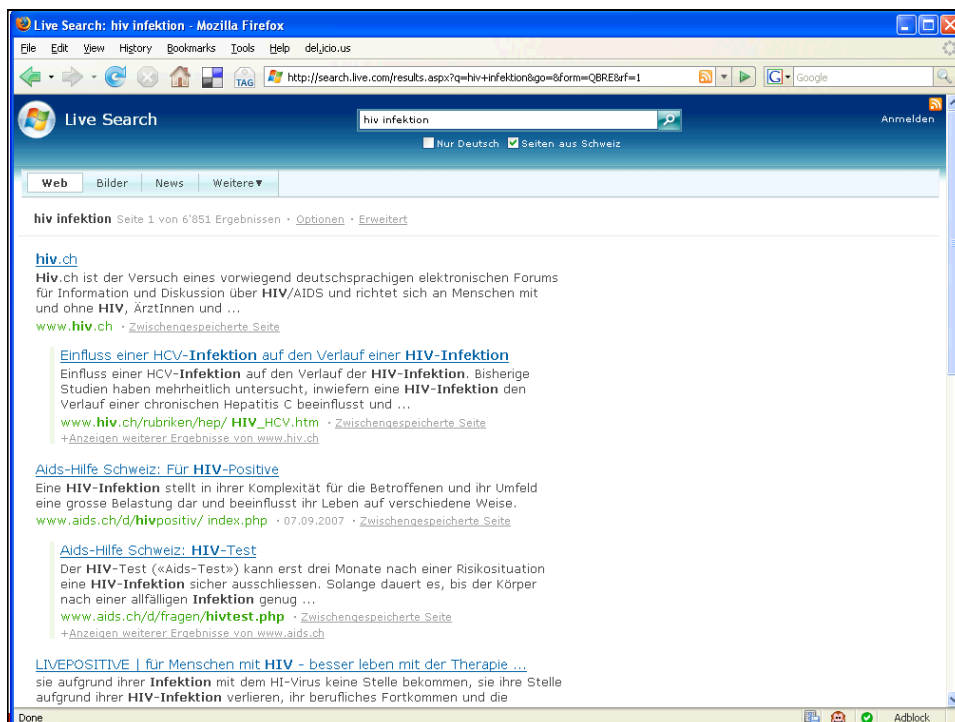
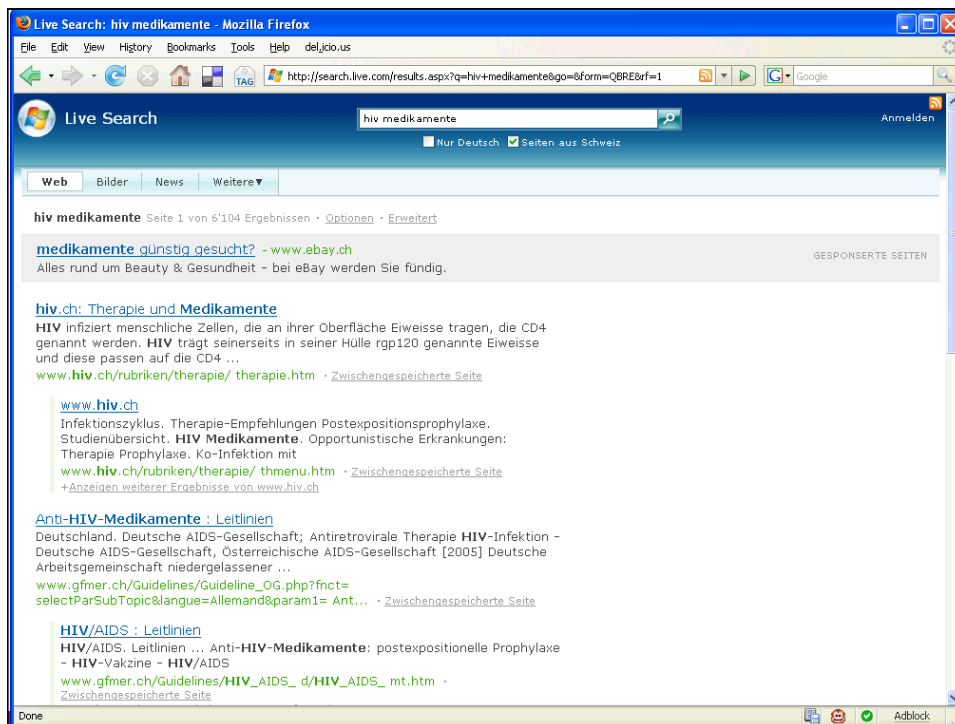
[www.msf.ch/Archiv-2006.629.0.html?&L= 1](#) · [Zwischengespeicherte Seite](#)

[Aids](#)

Prostratin ist nachweislich der erste Wirkstoff, der versteckte Viren im Körper entdeckt, die bisher gegen alle derzeitigen Aids-Medikamente resistent waren.

[www.habluetzel.com/aids.htm](#) · [Zwischengespeicherte Seite](#)

Done  AdBlock





## Unterschiedliche Daten – 1

» Einzelne Objekte haben kaum formale Attribute

» Typische Beispiele

- Von verschiedenen Usern gepflegte Webseiten (Internet)
- Nutzergenerierte Inhalte
- Wissensdomäne mit geringer Standardisierung

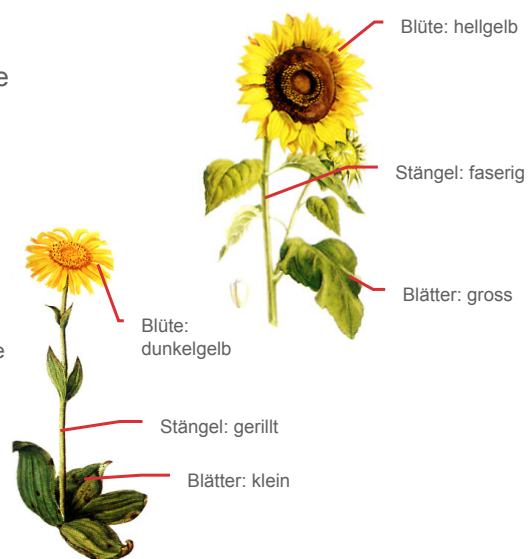


## Unterschiedliche Daten – 2

» Objekte haben formale Attribute und die Attribute haben eine bestimmte Ausprägung

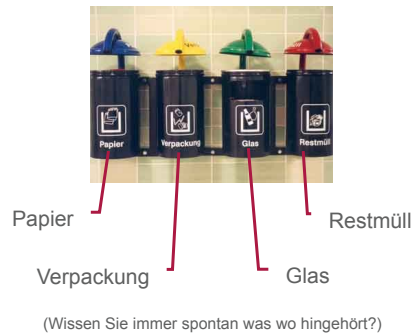
» Typische Beispiele

- Datenbanken: ERP, CRM...
- Telefonbuch
- Kleinanzeigen: Wohnung, Autos
- etc.



## Unterschiedliche Daten – 3

- » Mischformen
  - Formale Kriterien plus Inhalt
  - Nachträgliche Verschlagwortung (automatisiert oder manuell)
  - Anforderungen der Erschliessung
- » Typische Beispiele
  - Intranet (mehrdimensionaler Zugang)
  - Desktop / persönliche Dateien (Zeit plus „private“ Schlagworte)
  - Bücher in Bibliothek (Aufstellsystematik)



## Unterschiedliche Daten -- Lösungsansätze

Unstrukturierte Daten	Strukturierte Daten	Annotierte Daten
↓	↓	↓
<input type="text"/> <input type="button" value="Suchen"/>	Zimmer: 1.5 bis 4      Miete: unlim bis unlim Ort: legal      Größe (m2): <input type="button" value="Suchen"/>	Staatskanzlei      Grosse Rat Staatsrechnung des Kantons Bern      Regierungsrat Statistik      Finanzen Stellenmarkt des Kantons Bern Stellvertreterzentrale für Lehrkräfte Steuererklärung online
Volltextsuche	Parametersuche	Alphabetischer oder thematischer Index

## Wonach sucht der User?



## Oder was ist hier seine Sucheingabe?



Quelle: <http://www.google.com/jobs/britney.html>



Yahoo! Suchergebnisse für Läufer - Mozilla Firefox

File Edit View History Bookmarks Tools Help deljoo.us

http://ch.search.yahoo.com/search?ei=UTF-8&fr=sfp&p=L%C3%A4ufer&meta=vc%3Dcou

Web Bilder Video Audio Nachrichten

YAHOO! SUCHE Läufer

Suche

im Web Seiten auf Deutsch Seiten aus der Schweiz

Erweiterte Suche Einstellungen

Suchergebnisse 1 - 10 von ungefähr 566.000 für Läufer - 0.09 Sek (Réputat. en français) (Risultati in italiano) | Familienfilter ist AUS

SPONSOREN-LINKS

- Läufer - günstig bei BADER**  
www.bader.ch - Alles für und rund um die Themen Kleidung, Wohnen, Schlafen, Küche, Bad, Gesundheit, Körperpflege und viele weitere - grosse Auswahl, günstige Preise und kostenlose Lieferung frei Haus.
- Tischläufer nur CHF 14,95**  
www.tchibo.ch - Frisches Dessin: Tischläufer nur jetzt erhältlich bei Tchibo.
- Läufer bei ricardo.ch**  
www.ricardo.ch/wohnen - Bad, Sauna, Dekorationen, Designs, Esszimmer, Wohnzimmer, Kinderzimmer.

SPONSOREN-LINKS

**läeufner zu Traumpreisen**  
Möbel & Wohnen sicher kaufen. eBay Käuferschutz bis 1500.- Fr.  
www.ebay.ch

So werden Sie Sponsor...

- Läufer (Schach) - Wikipedia**  
Zu Beginn einer Partie befinden sich zwei weiße und zwei schwarze **Läufer** auf dem Brett. Jede Partei hat jeweils einen weißfeldrigen und einen schwarzfeldrigen **Läufer**. ...  
de.wikipedia.org/wiki/L%C3%A4ufer\_(Schach) - 45k - Im Cache - Weitere Seiten dieser Web-Site
- Läufer - Wikipedia**  
eine Figur im Schachspiel, siehe **Läufer** (Schach) ... eine Kurzform für **Laufers Schwein** (Hausschwein von 5–25kg) ... eines Hochbordsteines liegen **Gosseläufer** ...  
de.wikipedia.org/wiki/L%C3%A4ufer - 17k - Im Cache - Weitere Seiten dieser Web-Site
- Willkommen bei der Läuferelite March**  
Willkommen auf der Homepage der **Läuferelite March** ... Sonntag 06.30h Altendorf Dorfgeden (draussen) Long Jog © Copyright by **Läuferelite March** ...  
www.laeufer-elite-march.ch - 9k - Im Cache - Weitere Seiten dieser Web-Site
- läeufnergruppe-kerzers.ch - Home**  
**Laeufergruppe-Kerzers** - Die sympathische Seeländer Laufgemeinschaft ... standen unzählige Zuschauer und trugen die **Läufer** förmlich den Berg hinauf. ...  
www.laeufergruppe-kerzers.ch - 15k - Im Cache - Weitere Seiten dieser Web-Site

Done

namics

## Welchen Läufer meint der Suchende?

» Die nahe liegenden?

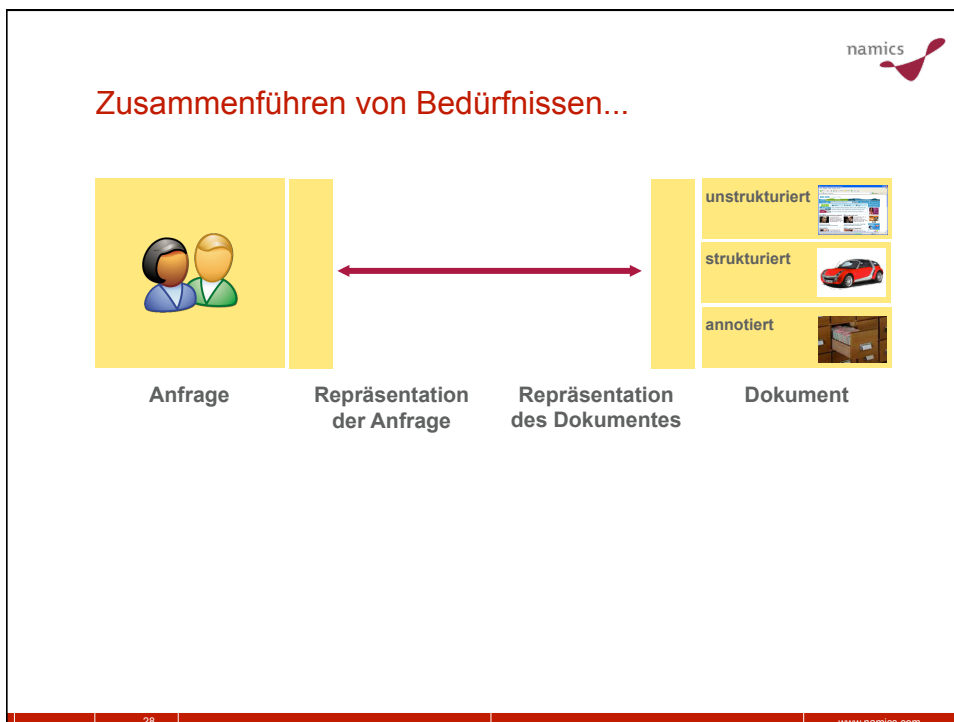




» Wikipedia kennt 24 Bedeutungen für "Läufer"

26

www.namics.com



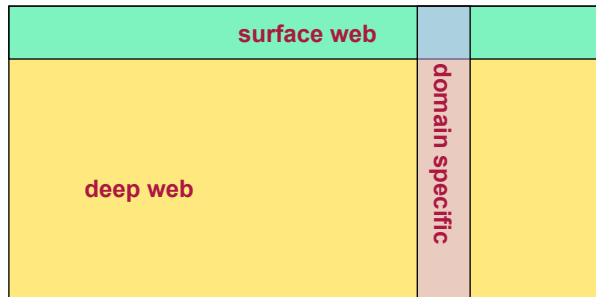
## Probleme der Zusammenführung

- » Meist ein Wortvergleich aber
  - wie drückt der User seinen Informationswunsch aus?
  - wie ist die Information in der Datenquelle vorhanden?
- » Synonyme: Handy <> Mobiltelefon <> Natel etc.
- » Unterschiedliche Schreibweisen
  - Jürg Stuker, Juerg Stuker, JÜRIG STUKER
  - J. Stuker, J. Stucker, Jürg St.
- » Produktive Sprachen
  - z. B. Fussballeuropameisterschaft
- » Mehrfach belegte Wörter
- » Unterschiedliche Sprachen
- » Unterschiedliche Repräsentation z. B. Bildsuche

## Plus: (öffentliche) Websuche versus Unternehmenssuche (1 von 2)

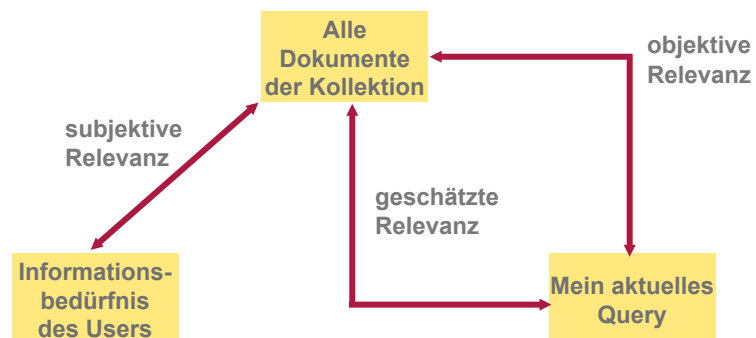
- » ~~Machen Sie die Suche genau so wie Google...~~
- » Nicht nur User, Daten, Suchaktivitäten und die finanzielle Investition sind in ihrem Intranet anders als bei Google (public) aber auch
  - Berechtigungssteuerung / Sicherheitsanforderungen
  - Verlinkung resp. Hierarchisierung der Seiten
  - Suchen auf eigenen Daten
    - Bedarf an Ausbeute (jede[r] will gefunden werden)
    - Aktualität ("Vanity Search")
  - Anwendungen (inkl. E-Mail, CRM, ERP etc.) und damit verbunden Strukturierung der Daten
  - Binäre Datenformate und lokale Dateien
  - Datenmenge
  - ...

## Plus: (öffentliche) Websuche versus Unternehmenssuche (2 von 2)



- » Unternehmenssuche ist Ausbeuteorientiert (ein bestimmtes Dokument resp. alle Dokumente zu einem Thema) häufig → Die lässt sich nur mit einem vollständigen (und aktuellen) Index erreichen.

## Plus: Relevanz ist subjektiv (Es gibt kein "perfektes" Retrievalergebnis)





Vorstellung des Produktportfolios

100% namics 

Eine Trefferliste mit unterschiedlichen  
Quellen und Berechtigungen

100% namics 

## Eine typisches Intranet-Suche



The screenshot shows a search interface with the following results:

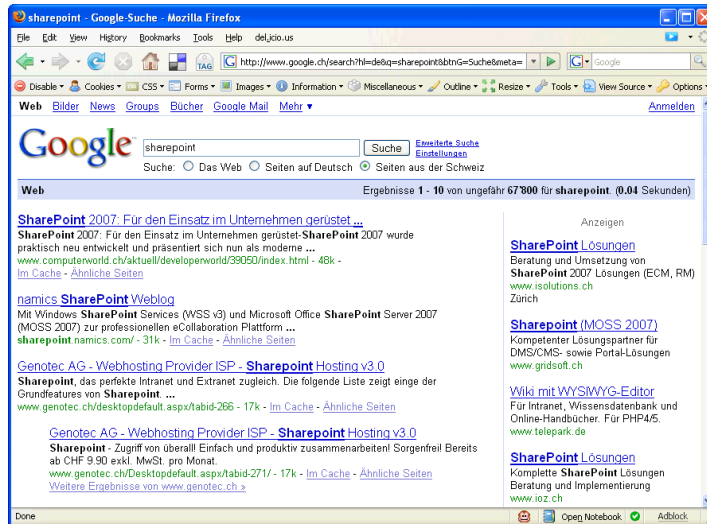
- Jürg Stuker**: CEO, Partner, Corporate, St. Gallen. Contact info: +41 71 228 6 701, +41 79 252 87 01. Email: jurg.stuker@namics.com.
- Antwort auf Request for Information**: E-Mail: jurg.stuker@namics.com. Alter: 35. Nationalität: CH. **Lebenslauf**, Studium der Wirtschaftsinformatik an der Universität Zürich, studienbegleitend ...
- Antwort auf Request for Information**: Deutsche Post World Net WCMS Customization Project **Lebenslauf** Projektleiter, namics (deutschland) sa 10. September 2003. Projektleitung namics ...
- Technologie Forum / Kurzinformation für die Moderation**: einen Kurz-**Lebenslauf** von sich zuzustellen ... mitarbeiterliste/%26sum%3Brg\_Stuker/ Mit freundlichen Grüßen Jürg "Geldinger Franziska" franziska
- CMS Evaluation**: ... Dadurch ist Ixos in der Lage den Kunden von der Entstehung einer Informationen, über den ganzen **Lebenslauf** bis hin zur endgültigen Aufbewahrung oder ...
- Ergänzt WAL-Check: Barrierefreie Webseiten**: ... 2004 namics sq Luzia Hüfen / Jürg Stuker edl, WAL\_Check\_14MAV04\_LH4\_02.doc Page 12 ... 3.2.3 PC **Lebenslauf** - www.edl.admin.ch 0% 20% 40% 60% 80% 100% Scanning ...

## Qualität der Rangliste

100% namics



## Alle einverstanden?



## Tipps bei der Einführung einer unternehmensweiten Suche

100% namics



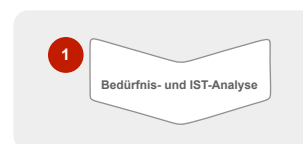
## Etwas Struktur kann nicht schaden



- » Prozess zur Auswahl und Einführung von Enterprise-Search-Produkten
- » Ausgehend von den wirklichen Problemen und Bedürfnissen Ihrer Nutzer
- » Produktunabhängige Gewichtung und Priorisierung
- » Aufzeigung der zugesicherten Eigenschaften
- » Objektiviertete Erfolgskontrolle

## 1a: Lernen Sie Ihre Anwender kennen

- » Nach **was** suchen die User
  - Zielgruppen
  - Informationsbedürfnisse
  - Häufige Anwendungsfälle
  - Aktuelle Problemstellungen
- » **Wie** suchen die User
  - Nutzertypen
- » Vorgehensweisen
  - Benutzer-Interviews
  - Benutzer-Beobachtung
  - Experten-Analysen
  - Nutzungs-Statistiken, Lernen im Betrieb



## 1b: Lernen Sie Ihre Quellen kennen

- » Vorhandene Daten und Anwendungen
- » Datenformate und -typen
- » Zugang resp. Schnittstelle
- » Netzwerkanbindung und Protokolle
- » Umfang und Aktualisierung
  
- » Vorgehensweisen
  - Inventarisierung
  - Analyse der integrierten Suchanwendung
  - Analyse von Schnittstellendefinitionen
  - Analyse von Testdatensätzen / Screens

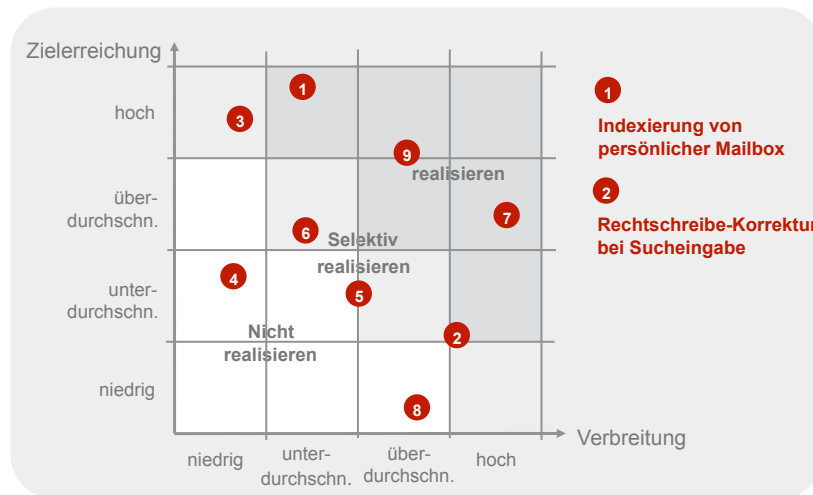


## 2: Anforderungspriorisierung

- » Anforderungen
  - Zusammenführung von Informationsbedürfnissen und der Antwortmöglichkeiten
  - Können und Wissen der Anwender
- » Rahmenbedingungen
  - Sicherheitsrichtlinien, Netzwerk-Architektur, Budget etc.
- » Priorisierung
  - Zielerreichung
  - Verbreitung: Anzahl Nutzer \* Nutzungshäufigkeit



## Beispiel: Priorisierungs-Methodik



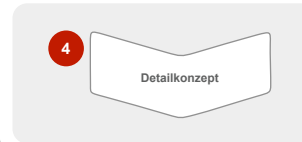
## 3: Grobkonzept und Proof of Concept

- » Grobkonzeption
  - Priorisierte Anforderungen
  - Lösungsansätze
  - Rahmenbedingungen
- » Produkt-Evaluation
  - Zusätzliche Faktoren wie Vendor Viability
  - Anzahl Anbieter vorselektieren
  - Produktpräsentationen / Schönheitswettbewerbe
- » Proof of Concept und Preisverhandlungen mit 1-2 Anbietern
  - Stärken & Schwächen Out-of-the-Box Features
  - Zeigt notwendige Aufwände für Implementierung



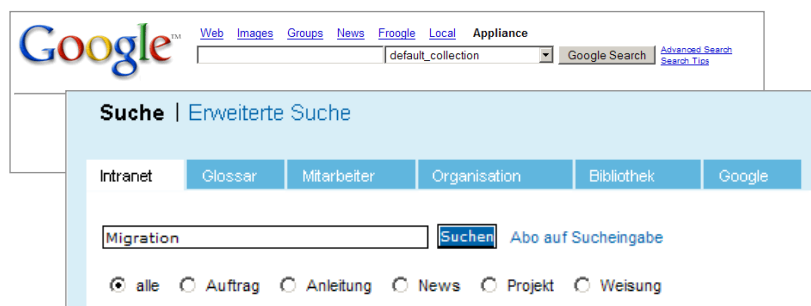
## 4: Konzeption der zu implementierenden Lösung

- » Frontend
  - Benutzerorientierte Interface-Gestaltung
  - Beschreibung der Funktionalitäten
  - Integration von Umsystemen (CMS, DMS)
- » Inhalte
  - Gruppierung von Inhalten / Kollektionen
  - Klassifizierung von Inhalten / Taxonomien
  - Zugriffsrechte und Benutzergruppen
- » Daten/Quellen
  - Schnittstellen, Integration von Umsystemen (CMS, DMS)
  - Indexierung (Vollständigkeit, Effizienz und Aktualität)
  - Berechtigungen
- » Technische Architektur



## 5: Technische Implementierung

- » Integration
- » Testing
- » Inbetriebnahme

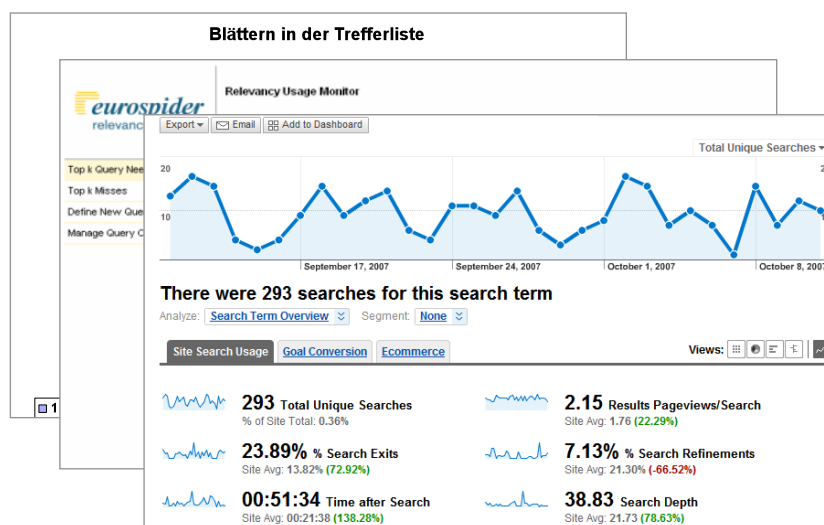


## 6: Erfolgskontrolle und iterative Optimierung

- » Feineinstellung des Systems im laufenden Betrieb
- » Fortwährende Verbesserungen (z.T. massiv)
- » Ausweisung des Erfolgs gegenüber Ausgangslage
- » Reevaluation der Bedürfnis- und IST-Analyse!
  
- » Vorgehensweise
  - Search Analytics



## Nutzungsanalyse und Tools für die Optimierung von Suchsystemen





## Tipps

- » Der Betrieb einer Suchlösung benötigt mehr Aufmerksamkeit als die ursprüngliche Implementierung
- » Sehr rasch wenige (wichtige) Fälle anbieten und dann ausbauen
- » Glauben Sie keinen Verkäufern („Don't throw money at technology...“)
- » Glauben Sie der Menge an Daten (“Statistik”)
  
- » NOGO: Visuell kreatives User Interface
- » NOGO: Suche als Sales-Tool

## Fragen & Antworten

100% namics 

## Anhang

100% namics 

## Weiterführende Informationen

- » Evaluation der Suchfunktion von CH Unternehmens-Websites
  - [http://blog.namics.com/2006/11/studie\\_evaluati.html](http://blog.namics.com/2006/11/studie_evaluati.html)
- » Search Analytics
  - <http://www.searchtools.com/slides/ess05/search-log-analysis.htm>
  - <http://www.rosenfeldmedia.com/books/searchanalytics/>
  - [http://blog.namics.com/2007/09/search\\_analytics.html](http://blog.namics.com/2007/09/search_analytics.html)
- » Klassiker der Informationssuche
  - Daniel E. Rose, Danny Levinson. Understanding User Goals in Web Search.
  - Jennifer English. Enterprise Search: A View From the Trenches.
  - Ellen M. Voorhees. The Philosophy of Information Retrieval Evaluation.
  - Jan Pedersen. The Four Dimensions of Search Engine Quality.
  - Maarten de Rijke. Introduction to Information Retrieval, Beyond Document Retrieval.
- » [http://blog.namics.com/information\\_retrieval/](http://blog.namics.com/information_retrieval/)

## Eckdaten zu namics

- » Marktführender Schweizer Berater für Internet-Anwendungen und E-Business, starke Präsenz in Deutschland
- » Fokus
  - Strategieberatung für Internet
  - Konzeption und Implementierung nutzergerechter, effizienter und begeisternder Internet-Anwendungen
  - Vermarktung und Lancierung von Online Aktivitäten
  - werblich orientierte Markenkommunikation
- » Zahlen und Fakten
  - gegründet 1995 als Spinoff der Universität St. Gallen
  - 240 Mitarbeiter, Umsatz 2007 CHF 34,2 Mio.
  - Standorte: Bern, Frankfurt, Hamburg, München, St. Gallen, Zug, Zürich
  - PubliGroupe SA, Minderheit beim Management (15 Partner)

## Eine Auswahl unserer Kunden



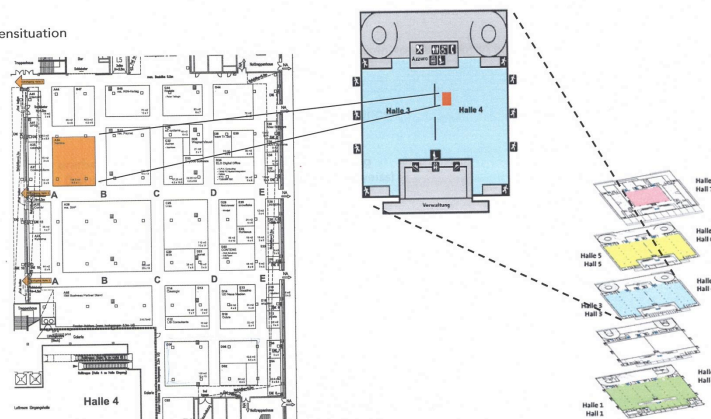
## Alle namics Referate auf eine Blick

- » 20.05. 09.15-10.45 Top 10 Internet-Standards der Zukunft
- » 20.05. 13.15-14.45 Flash vs. Silverlight: Das grosse Duell
- » 21.05. 09.15-10.45 Können Software-Innovationsprojekte gefördert, geplant und geführt werden?
- » 21.05. 11.15-12.45 CMS-Shoot-out: Content-Management-Systeme im Vergleich
- » 21.05. 13.15-14.45 Internet für alle – Virtuelle Grenzen überwinden
- » 22.05. 09.15-10.45 Web Analytics in der Praxis
- » 22.05. 11.15-12.45 Effiziente Rezepte für die Online-Kundengewinnung
- » 22.05. 15.15-16.45 CMS-Migration Best Practices
- » 23.05. 11.15-12.45 Enterprise Search: Google vs. Microsoft
- » 23.05. 13.15-14.45 Weblogs und Wikis im Firmeneinsatz
- » 23.05. 15.15-16.45 TV 2.0 im Business-Einsatz
  
- » Die Handouts stehen als Download zur Verfügung: [www.namics.com/knowledge](http://www.namics.com/knowledge)

## Besuchen Sie uns an der Orbit-iEX

- » namics Stand in der Halle 4/A34

Hallensituation



Besten Dank für Ihre Aufmerksamkeit!



René Knöfel, [rknoefel@microsoft.com](mailto:rknoefel@microsoft.com)  
Joachim Glaubrecht, [joachim@google.com](mailto:joachim@google.com)  
Marco Billeter, [marco.billeter@fastsearch.com](mailto:marco.billeter@fastsearch.com)  
Jürg Stuker, [juerg.stuker@namics.com](mailto:juerg.stuker@namics.com)

Bern, Frankfurt, Hamburg, München, St. Gallen, Zug, Zürich