



Zürcher  
Hochschule  
Winterthur

InIT

# Evaluation der Suchfunktion von Schweizer Unternehmens-Websites: Gesamtresultat

Dr. Martin Braschler

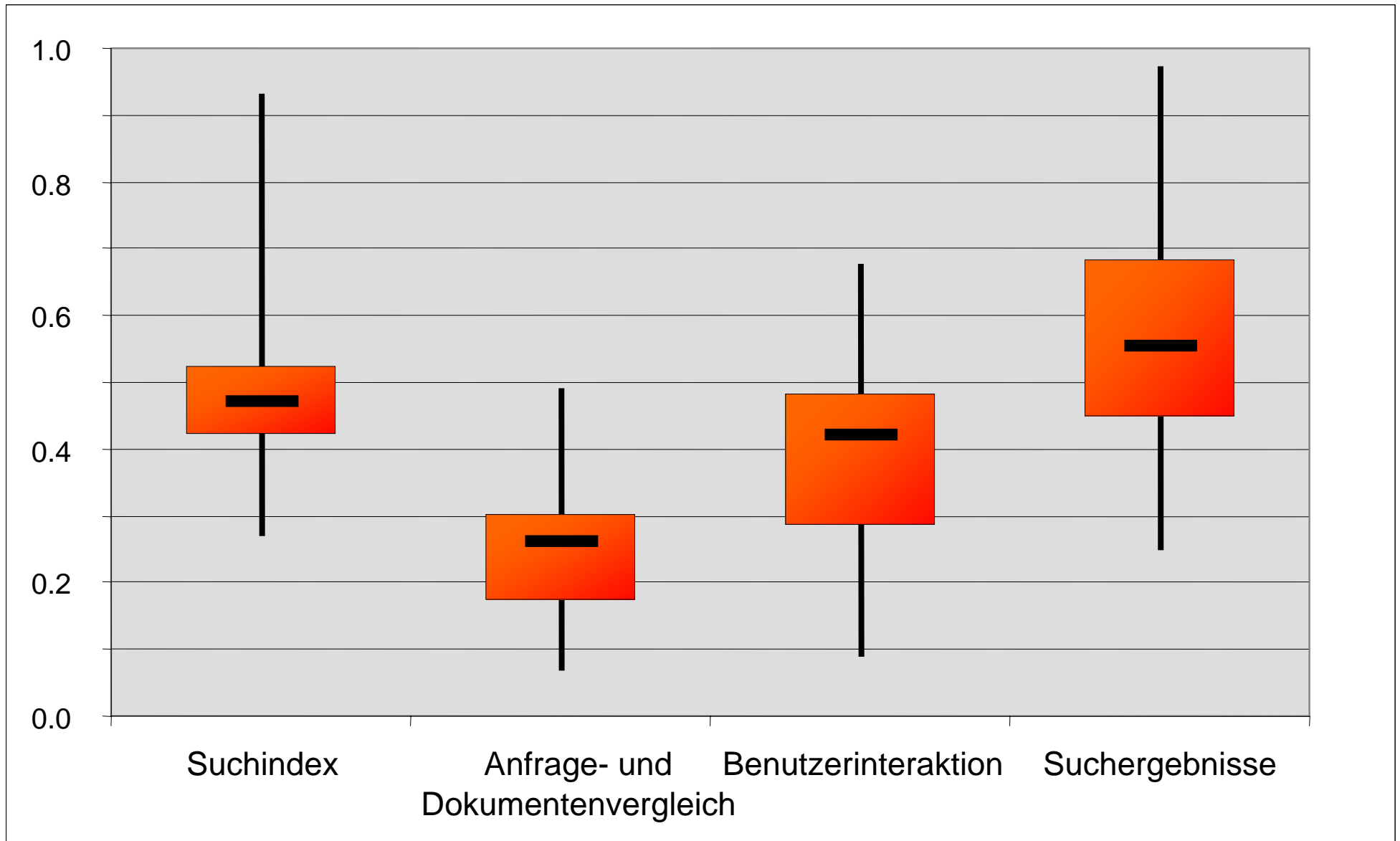
Dozent, Institut für angewandte Informationstechnologie  
Zürcher Hochschule Winterthur

30.11.2006

# Inhalt

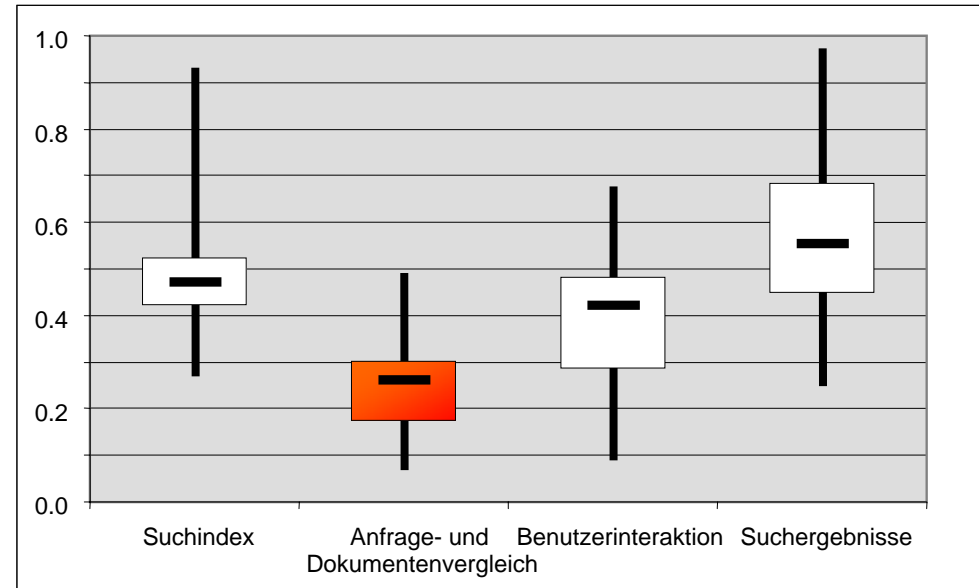
- Gesamtergebnisse
- Analyse
- Allgemeiner Eindruck/Spasfaktor
- Empfehlungen

# Gesamtresultate Hauptkriterien



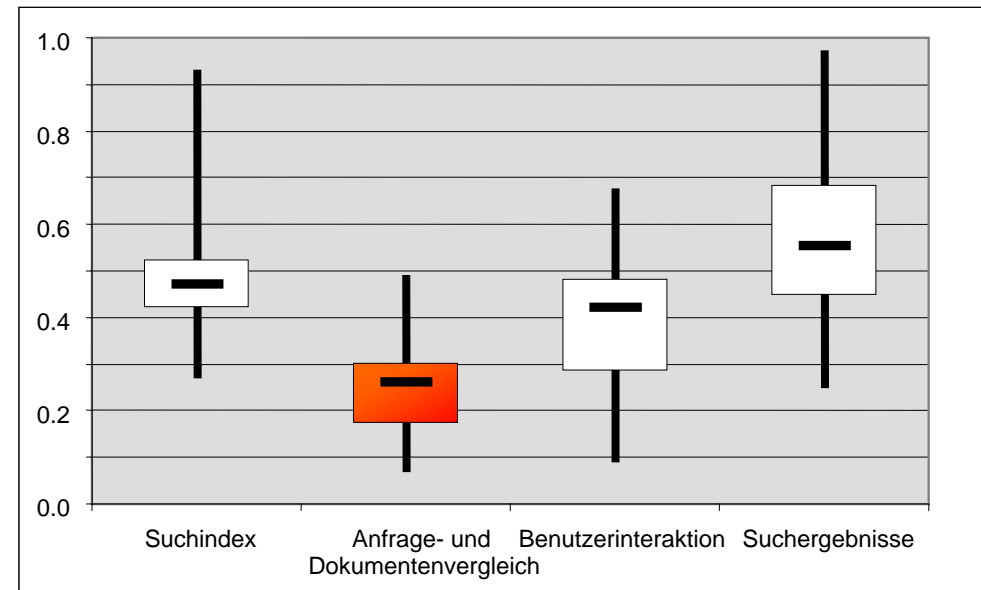
# Absolute Leistungswerte

- Absolute Leistungswerte sind grundsätzlich vorsichtig einzuordnen
- Die Leistungswerte im Bereich "Anfrage- und Dokumentenvergleich" fallen durchs Band gegenüber den anderen Kriterien ab (Median 0.26, andere Kriterien  $>0.4$ , niedrigstes Minimum, niedrigstes Maximum)
- Die Streuung dabei am zweitgeringsten  $\rightarrow$  die Leistung der grossen Mehrheit der Site-Suchen in diesem Bereich muss als "ungenügend" betrachtet werden



# Absolute Leistungswerte

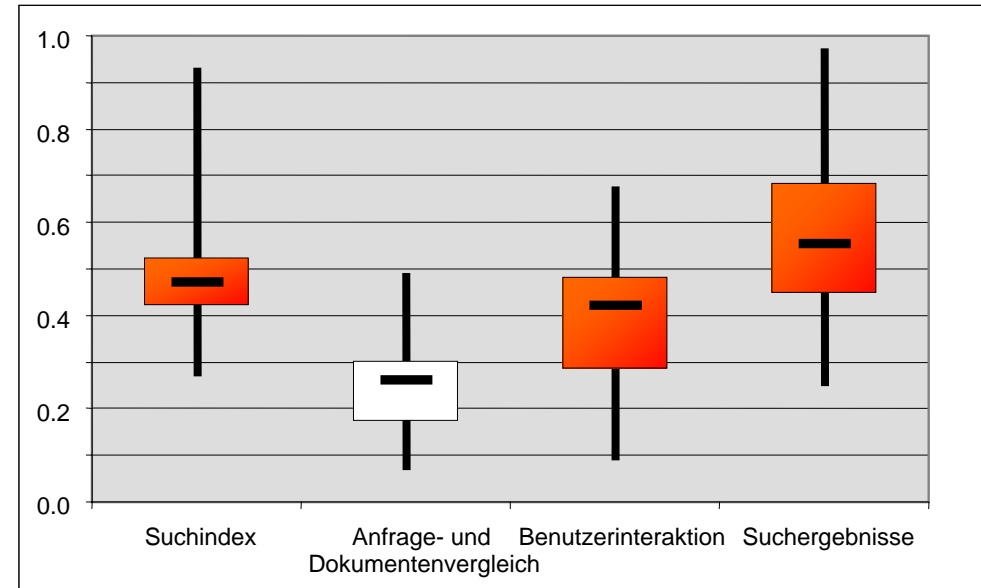
- Auch die maximal erreichte Leistung für Anfrage- und Dokumentenvergleich ist sehr tief.



- Nur gerade drei weitere Websites erreichen mindestens 80% der Leistung dieses "Klassenbesten"

# Streuung

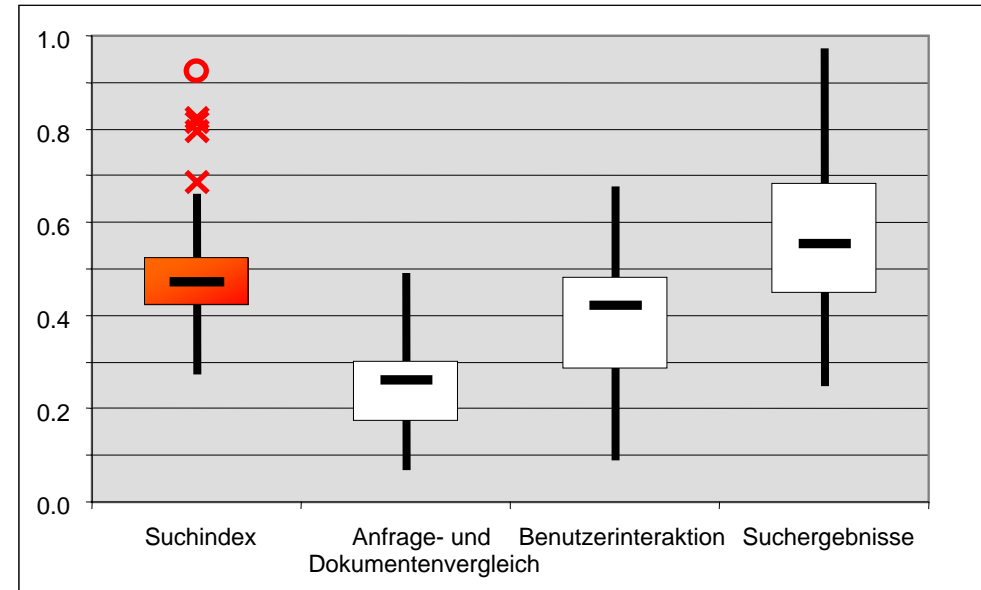
- Die restlichen drei Hauptkriterien befinden sich in der mittleren Leistung um 0.5
- Die Streuung ist aber stark unterschiedlich (siehe Interquartilsabstand)



Statistiken	<b>Suchindex</b>	A&D- Vergleich	<b>Benutzer- interaktion</b>	<b>Suchergeb.</b>
Maximum	0.93	0.49	0.68	0.97
Median	<b>0.47</b>	0.26	<b>0.42</b>	<b>0.55</b>
Minimum	0.27	0.07	0.09	0.25
IQ-Abstand	<b>0.10</b>	0.13	<b>0.20</b>	<b>0.23</b>

# Schiefe

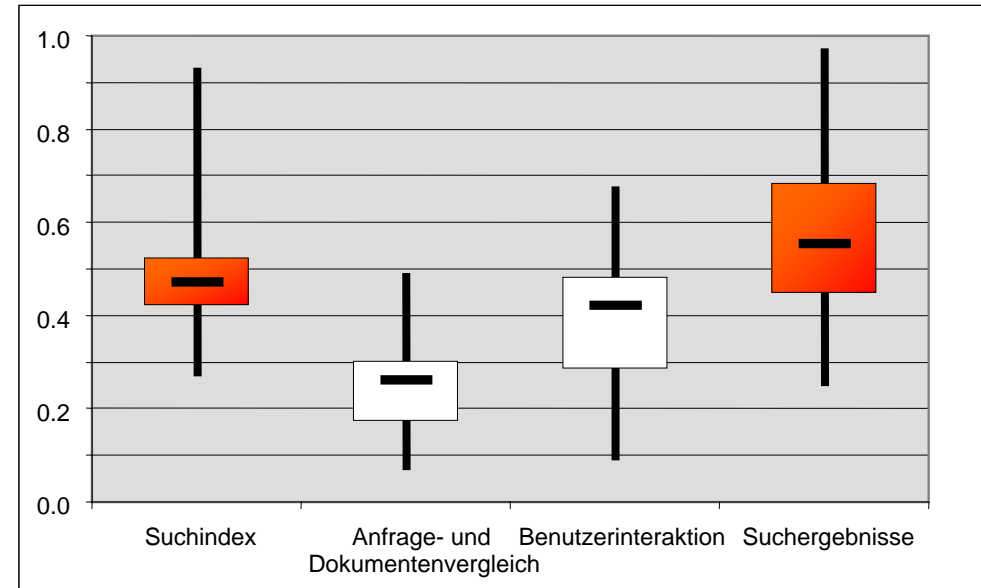
- Die Lage des Medians in der Box ist für das Kriterium "Suchindex" bemerkenswert symmetrisch



- Anders sieht es an den beiden Enden der Leistungsskala aus: der Vorsprung der besten Probanden auf das "Hauptfeld" ist mehr als doppelt so gross wie auf die "Nachzügler"
- Es existieren Ausreisser nach oben; drei "schwache" Ausreisser und ein "extremer" Ausreisser

# Bestleistungen

- Für die beiden Kriterien "Suchindex" und "Suchergebnisse" gibt es Sites, die nahe an das theoretische Maximum kommen



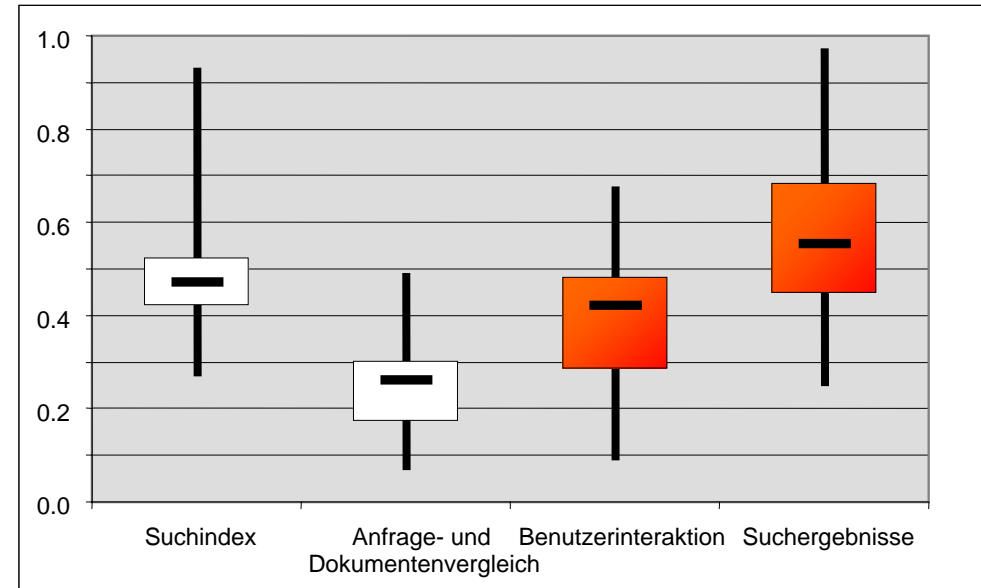
- Es handelt sich dabei aber für beide Kriterien um Einzelfälle
- Auch diese Probanden waren nicht ohne Fehl und Tadel

Statistiken	<b>Suchindex</b>	A&D- Vergleich	Benutzer- interaktion	<b>Suchergeb.</b>
Maximum	<b>0.93</b>	0.49	0.68	<b>0.97</b>



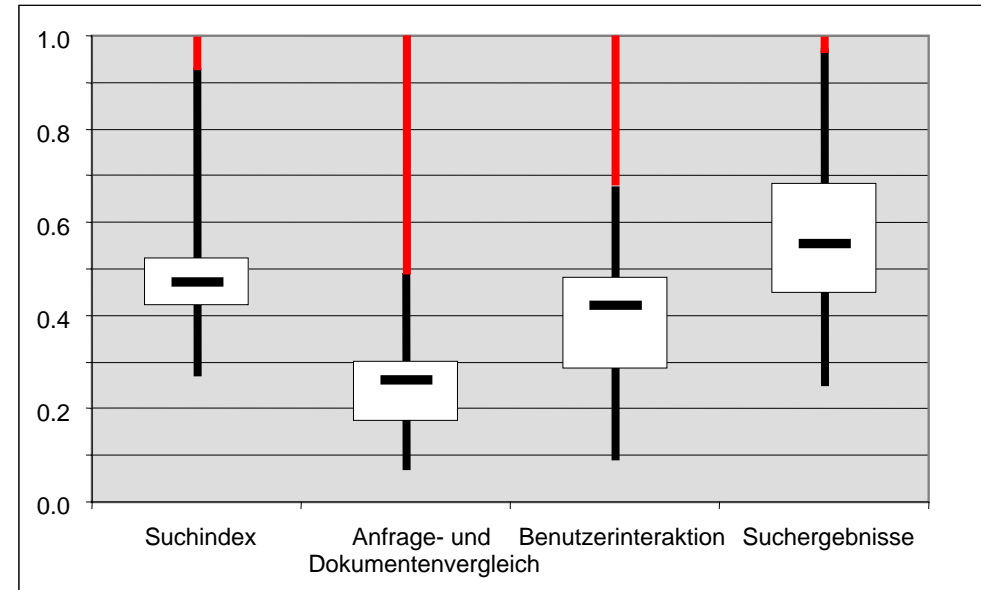
# Bandbreite

- Die grösste Bandbreite ist zu verzeichnen für das Kriterium "Suchergebnisse"
- Der Vorsprung der "Klassenbesten" ist grösser als der Rückstand der "Nachzügler"
- Am meisten "Nachzügler" existieren für das Kriterium "Benutzerinteraktion". Hier ist der Abstand der schwächsten Probanden auf die mittlere Leistung am grössten.



# Luft nach oben

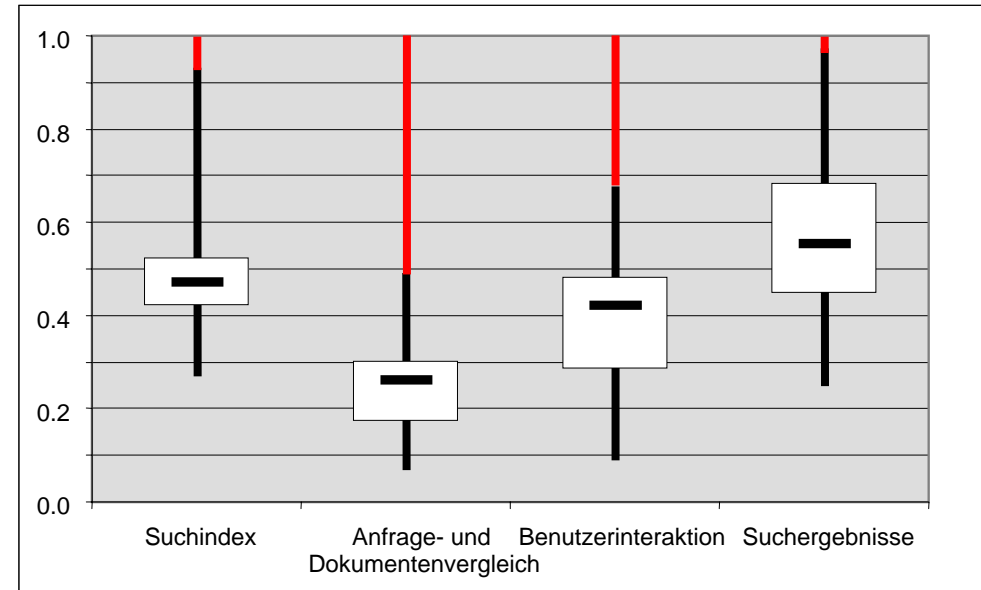
- Es besteht nicht nur eine teilweise grosse Lücke zwischen den erreichten Maxima und dem theoretischen Höchstwert



- Die vier Bestleistungen wurden von vier unterschiedlichen Site-Suchen erreicht

# Luft nach oben

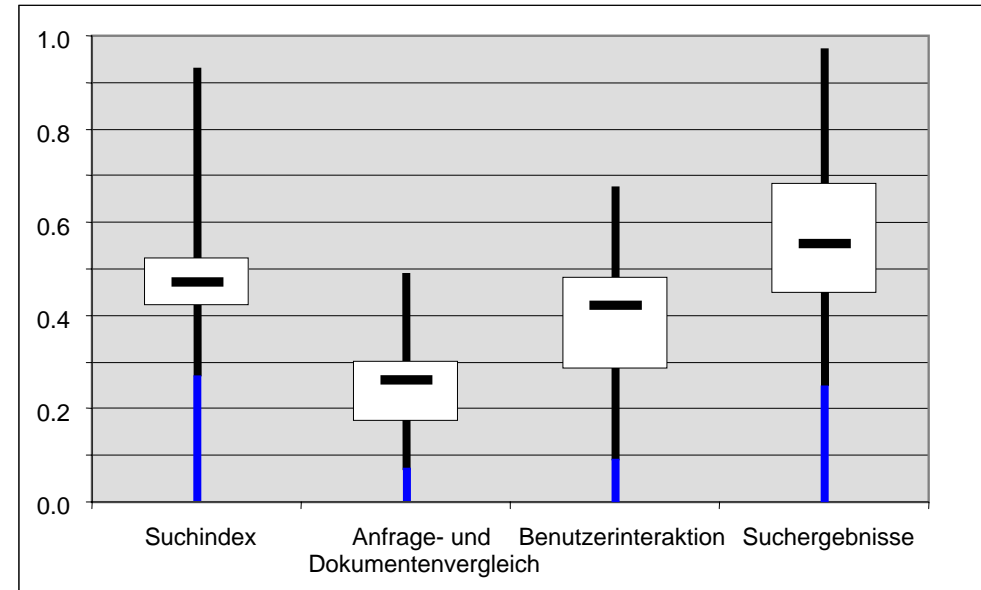
- Betrachtet man engere Spitzengruppen, so erreicht kein Proband in allen Kategorien Bestnoten



- Nur gerade drei Probanden sind konsistent in der erweiterten Spitzengruppe für alle Kriterien
- Die gleichen Probanden sind für jeweils drei Hauptkriterien unter den Allerbesten (engere Spitzengruppen)
- Mehr als die Hälfte der Probanden konnte sich für kein Hauptkriterium in der engeren Spitzengruppe etablieren

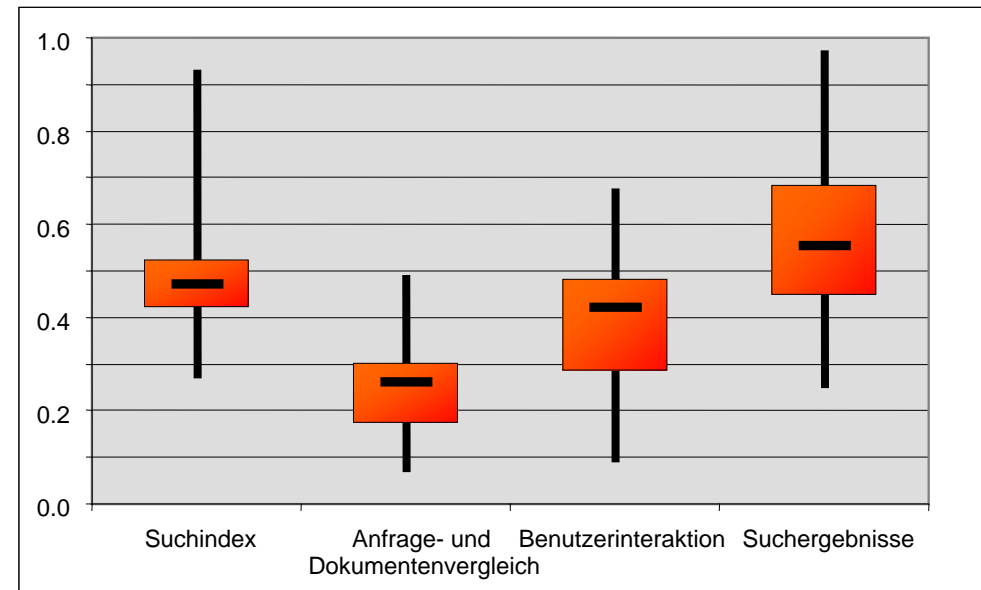
# Nachzügler

- Kein Proband ist für alle Hauptkriterien in der untersten Gruppe gelandet
- Ein Proband ist für drei der vier Kriterien in der schwächsten Gruppe angesiedelt
- Fast die Hälfte der Probanden ist für mindestens ein Hauptkriterium bei den schwächsten Leistungen anzutreffen



# Paarweise Betrachtung

- Betrachtet man die Leistungen paarweise zwischen Hauptkriterien, so zeigt sich kein einheitliches Bild



- Für kein Paar von Hauptkriterien zeigen konsistent dieselben Probanden gute Leistungen
- Dies deutet darauf hin, dass die Kriterien gut gewählt wurden (weitgehend orthogonal)

# Allgemeiner Eindruck

"Soft Questions":

- Gesamteindruck (Note 1/ungenügend..10/exzellente)
- Spassfaktor/Sonstige Bemerkungen (Auffälligkeiten, die nicht durch die Kriterien der Studie abgedeckt sind)

# Allgemeiner Eindruck

Gesamteindruck:

- Durchschnittsnote: 4.84
- Median: 5
- Deckt sich fast schon gespenstisch gut mit den absoluten Medianwerten dreier der vier Hauptkriterien

# Allgemeiner Eindruck

Gesamteindruck: (Kommentare)

- Tester vermissten immer wieder "griffige" Einschränkungsmöglichkeiten
- Es wurde vermerkt, dass oft Operatoren fehlen/hart verdrahtet sind, um auf die Suche Einfluss zu nehmen (vermisst: Phrasenoperator)
- Mehrmals Erstaunen Ausdruck gegeben, dass die Suche so "simpel" ausgeführt ist
- Einige "blumige" Erkenntnisse ("die Darstellung der Treffer ist kriminell") 😊



# Spassfaktor

## Spassfaktor/Sonstige Bemerkungen

- 0 = machte keinen Spass ☹️
  - 1 = war okay 😐
  - 2 = machte viel Spass 😊
- 
- Durchschnitt: 0.82
  - Median: 1

# Spassfaktor

Kommentare:

- **Verdarb den Spass:**
  - Fehleranfällig
  - Langsam
  - Unerwartetes Verhalten (Bsp.: nur Kleinschreibung funktioniert)
  - Vorgehensweise nicht transparent
  
- **Machte Laune:**
  - einfach
  - schnell
  - ansprechende Präsentation

# Empfehlungen

- Die Studie hat viel gute Beispiele für die erfolgreiche Implementation einer Site-Suche aufgezeigt
- Gleichzeitig wurde aber auch ein grosses Potential „nach oben“ aufgedeckt
- Die folgenden Empfehlungen umfassen Punkte, die einen besonders grossen Anteil an den Streuungen in den Hauptkriterien haben
- Diese Punkte haben sich auch im Urteil der Tester als besonders wichtig ausgezeichnet

# Empfehlungen

- Mehr Gewicht auf Vollständigkeit und Aktualität des Suchindex
  - ➔ Die Suche ist nur so gut wie die Datenbasis
  - Viele Websites haben Verbesserungsbedarf
  - Aktives Einbinden der Suchfunktion in die Gesamtwebsite statt "passives Crawlen"
  
- Bessere Pflege der Metadaten
  - Garantieren einer Minimalqualität
  - Korrektes Datum, korrekter Titel, gute Trefferzitate
  - Erlaubt bessere Sucheinschränkungen

# Empfehlungen

- Orientierung an bewährten Standards für Benutzerinteraktion
  - Die Tester haben irritiert reagiert, wenn die Suchfunktion sich unerwartet oder intransparent verhalten hat.
  - Gegebenenfalls Orientierung an "horizontalen" Suchdiensten
- Ausbau der Qualität des Vergleichs zwischen Anfragen und Dokumenten
  - Behandlung natürlicher Sprache: Flexionen, Komposita
  - Mehrsprachigkeit
  - Wichtig: Nutzer sind auch ausbeuteorientiert!

# Empfehlungen

- Trefferliste nicht nur aus Dokumenten- sondern auch aus Nutzersicht aufbereiten
  - Die Website dient als Kommunikationsinstrument
  - Häufige Informationsbedürfnisse vorwegnehmen
  - Stratifizierung der Resultate, Entitätenerkennung, ..
- Der Suchfunktion die Bedeutung geben, die sie verdient
  - Wenn die Website beim Browsen zu langsam wäre, wäre sie schon lange "geflickt" worden
  - Warum nicht die Suchfunktion?