

Thomas Mandl, Teresa Märtil, Christa Womser-Hacker

Informationswissenschaft
Universität Hildesheim
mandl@uni-hildesheim.de

Mehrsprachige Suche in Social Tagging-Systemen

Eine Untersuchung am Beispiel von Flickr

Überblick



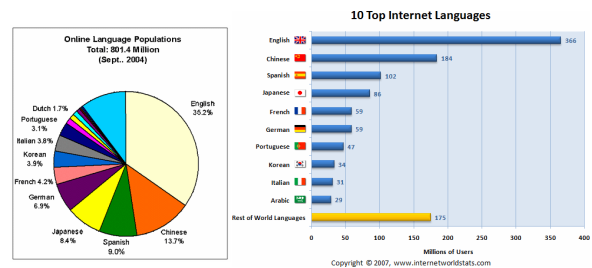
- Einleitung
- Flickr
- Fragestellung: Mehrsprachigkeit in sozialen Netzwerken
- Vorgehen für Experiment
- Ergebnisse
- Weitere Aspekte der Internationalität

Einleitung



- **Social Tagging-Systeme:** Nutzer versehen Ressourcen mit frei gewählten Schlagworten – **Tags**
- Ressourcen für sich organisieren und zugleich mit anderen teilen
- Delicious, Bibsonomy, Flickr, YouTube, MyBib ...

Sprachverteilung im Internet



<http://www.internetworldstats.com>

Wie multilingual ist das Web?

Dokumente in zahlreichen Sprachen
aber ...

- gibt es Links zwischen Dokumenten verschiedener Sprache?
- werden diese benutzt? (Theilwall 2005)
- gibt Dokumente mit mehreren Sprachen?
- was sind die Gründe für Mehrsprachigkeit? (Artemenko et al. 2007)

Mehrsprachiges Retrieval (Mandl @ CLEF 2002-2009)

- Web Usage Mining und Sprachidentifikation
- Rolle der geographischen Herkunft und der Kompetenz (Kralisch & Mandl 2006)

Einleitung

- **Internationale Nutzerbasis** vieler Web 2.0 Systeme
- Ressourcen werden in unterschiedlichen Sprachen getaggt
- Sprachendschungel, Ressourcen nicht universell zugänglich und auffindbar
- Welche Suchstrategien verwenden Benutzer angesichts der Mehrsprachigkeit?
 - Sind sie sich dieser bewusst?
 - Suchen sie mehrsprachig?

- Soziales Netzwerk zum Austausch von Fotos und Videos
- Organisation der Bilder durch Tags
- Zahlreiche Möglichkeiten zum Browsen
- Bilder werden vor allem dadurch und über das soziale Netzwerk gefunden

- Mehrsprachigkeit auf Flickr und Folgen für Suche nach Bildern:
 - Stört es die Nutzer, dass nicht alle Bilder über eine Sprache verfügbar sind?
 - Führen sie ihre Suche in mehreren Sprachen durch?
 - Welche Faktoren beeinflussen die Sprachwahl?

- Untersuchung angelehnt an **iCLEF** (*interactive* CLEF-Track)
 - Untersuchung von sprachübergreifendem Retrieval aus Nutzerperspektive
 - Erstmals 2006, 2008 mit eigens erstelltem Interface *Flickling*
 - Kritik: Flickr wird als reine Suchmaschine betrachtet, ist jedoch v.a. soziales Netzwerk

- Benutzertest mit 21 Teilnehmern:
 - 2 Aufgaben Taggen: Taggen eines Bildes + Bewerten vorhandener Tags
 - 3 Aufgaben Suchen: Vorgegebenes Bild suchen, Illustrieren verschiedener Schlagzeilen + freie Suche zu bestimmtem Thema



Source: Flickr

- Bewusster Umgang mit Mehrsprachigkeit
 - Englisch wird als wichtigste Sprache gesehen
 - Trotzdem werden Bilder auch in anderen Sprachen gesucht
 - Annahmen über Inhalte
 - Problem mangelnde aktive Kompetenz
 - Verständnis, dass Bilder nicht in allen Sprachen verfügbar

Ergebnisse 1/3

Suchsprache	Anzahl
Nur Deutsch	13
Deutsch + Englisch	8

Tabelle 1: Suchsprachen deutsche Schlagzeile

Suchsprache	Anzahl
Nur Englisch	13
Deutsch + Englisch	8

Tabelle 2: Suchsprachen englische Schlagzeile

Suchsprache	Anzahl
Nur Französisch	7
Französisch + Englisch	2
Französisch + Deutsch	1
Deutsch + Englisch	1

Tabelle 3: Suchsprachen französische Schlagzeile

Suchsprache	Anzahl
Nur Spanisch	1
Nur Englisch	5
Spanisch + Englisch	2
Spanisch + Deutsch + Englisch	1
Deutsch + Englisch	1

Tabelle 4: Suchsprachen spanische Schlagzeile

Ergebnisse 2/3

- Vor allem Englisch bei Suche verwendet
- Je nach Thema auch andere Sprachen
- Teilnehmer würden auch in unbekanntem Sprachen suchen
→ visuelles Feedback
- Suchverhalten: Versuch, Suchergebnisse über längere Suchanfragen zu spezifizieren
→ gegenteiliger Effekt

Ergebnisse 3/3

- Qualität der Suchergebnisse abhängig von Sprache, Suchbegriffe und Länge der Anfrage
→ beste Ergebnisse durch einsprachige, allgemeine und kurze Anfragen
- Spezifikation der Ergebnisse am besten durch Browsen in Tags, Cluster, Gruppen etc.

Fazit

- Teilnehmer sind sich der Problematik der Mehrsprachigkeit bewusst
- Sinnhaftigkeit von integrierten Übersetzungsmöglichkeiten in weiteren Studien untersuchen
→ vielfältige Wege zum Browsen und Finden von Bildern auf Flickr erhalten
- Integration von Übersetzungstools

Weitere Aspekte der Internationalität

- Applikationen im Social Web
 - Soziale Netze für weltweite Nutzung
 - Einfluss der lokalen Kultur auf Nutzung sehr stark
 - Empirische Studie
 - Soziales Kapital als Hauptgrund für die Nutzung von Netzwerken
 - Interpretation und Konstruktion von sozialem Kapital international sehr unterschiedlich
- (Mandl & Womser-Hacker @ ACM CHI WS 2008)
- > Quelle für erhebliche interkulturelle Mißverständnisse